MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 109 No. 5

MEETING OF 1949 OCTOBER 14

Professor W. M. Smart, President, in the Chair

The election by the Council of the following Fellows was duly confirmed :—

Noel Charles Clements Barrell, Colline Cottage, Shamley Green, Guildford, Surrey (proposed by H. H. Plaskett);

Ronald William Boggis, 41 Maycross Avenue, Morden, Surrey (proposed by F. M. Holborn);

Miss Yi-Hsin Chou, Sun Yat Sen University, Canton, China (proposed by F. J. M. Stratton);

George Forman Hutchinson, 53 Captain's Road, Edinburgh, 9 (proposed by N. G. Matthews);

Theodore McKenzie, 34 Market Street, Grahamstown, South Africa (proposed by R. H. Stoy);

Patrick Alfred Moore, Glencathara, Worsted Lane, East Grinstead, Surrey (proposed by D. W. G. Arthur); and

Wilfred Shaw, F.R.C.S., F.R.C.O.G., 109 Harley Street, London, W.1 (proposed by W. M. H. Greaves).

The election by the Council of the following Junior Members was duly confirmed :—

Alan Henry Batten, St. Mark's Vicarage, Chapel Ash, Wolverhampton, Staffordshire (proposed by C. D. Grimwade); and

Bertrand John Harris, 27 Ruthin Road, Blackheath, London, S.E.3 (proposed by A. Hunter).

Two hundred and sixty presents were announced as having been received since the last meeting, including :—

Table of Logarithms giving 5*-figure Logarithms of the numbers* 10,000–40,000 *and* 4,000–10,000 (presented by D. H. Sadler);

J. F. Cox, *Éléments d'Astronomie et de Géodésie* (presented by the author);

C. Hoffmeister, *Meteorströme* (presented by the author);

C. Walén, *On the Vibratory Rotation of the Sun* (presented by the author); and

G. J. Whitrow, *The Structure of the Universe* (presented by the author).

Professor Otto Struve delivered the George Darwin Lecture (see p. 487), taking as his subject " Spectroscopic Binaries ".

In introducing the lecturer the President said that "five years ago Professor Struve was awarded the Society's Gold Medal. At the time the war was still raging in three continents and all over the seven seas, and it is not surprising that Professor Struve found it impossible to be with us on that occasion. We, to-day, are the gainers and we give him a very warm welcome in a room with which he is very familiar.

" The name of Struve has a long and continuous association in astronomy. Wilhelm Struve, born nearly 160 years ago, was Director of Dorpat Observatory at a very early age, and in 1835 was invited by the Emperor Nicholas of Russia to supervise the erection of a great Observatory (Pulkova) on the outskirts of St. Petersburg, and he was its Director for 19 years. In 1826, he was awarded the Gold Medal of the Society for the discovery and measurement of double stars, a new branch of astronomy founded a little before by Sir William Herschel. Wilhelm, it may be added, was one of the three astronomers who, almost simultaneously, succeeded for the first time in measuring the vast distances of the stars.

"Wilhelm was succeeded at Pulkova by his son, Otto Struve, who was awarded the Gold Medal in 1850 for the determination of the constant of precession, taking account of the proper motion of the Solar System, or as we should now say, the Solar Motion. Hermann Struve, the elder son of Otto, founded the Observatory of Berlin-Babelsberg in 1913; ten years previously he had been awarded our Gold Medal for his investigations on the satellites of Saturn. Ludwig Struve, the second son of Otto, and father of our Lecturer, was Professor of Astronomy at Kharkov, but by some mischance failed to maintain the family tradition in the matter of Gold Medals. His son, whom we welcome here to-day, restored the continuity, the grounds of the award being his ' Observations and Interpretations of the Spectra of Stars and Nebulae '.

"Our Lecturer has been at Yerkes since 1922, and in 1932 he succeeded Frost as Director of that great institution. He founded the McDonald Observatory on Mount Locke in Texas, which was opened in 1939. His first paper was on the star 9 Argus, a visual double of period 23 years, observable also as a spectroscopic binary; in his brilliant address on the award of the Gold Medal to our Lecturer in 1944 April, Professor Milne remarked that ' the spirit of his great-grandfather was stirring within him '. In this address you will find a full account of Professor Struve's distinguished researches. Perhaps I ought to add that in recent months Professor Struve has been engaged in bringing new ideas, based on observations, into the perhaps most fascinating subject of all, ' The Evolution of the Stars ', about which we shall be able to read in a forthcoming book.

" It is interesting and significant that our Lecturer should return to his first interest, Spectroscopic Binaries, and without further delay I now invite him to deliver the George Darwin Lecture on this subject."

SPECTROSCOPIC BINARIES

George Darwin Lecture, delivered by Professor Otto Struve on 1949 *October* 14

Both Sir George Darwin, whose name is commemorated in these annual lectures, and Sir James Jeans, whose generosity made them possible, were animated throughout their lives by an overwhelming desire to explain the origin and the evolution of double stars and thereby to interpret the variety of observational results that had accumulated since Sir William Herschel recognized the physical connection between the components of most visual binaries. It is appropriate that a Darwin Lecture should be devoted to the problem of the closer binaries, those that can be observed with the help of the Doppler effect, because in them we must seek for clues to explain the origin and evolution not only of the binary stars themselves, but also of the rapidly-rotating single stars, and perhaps of all cosmic formations possessing large amounts of rotational momentum.

My own work on spectroscopic binaries started about a quarter of a century ago, and my first communication to the Royal Astronomical Society, in 1925 December, was entitled "A Study of the Nature of Spectroscopic Binaries". I should like to record that the kind and encouraging words spoken by the President, Sir James Jeans, in connection with this paper served as a real inspiration to me in later work, and it is perhaps due to this circumstance that I am able to review the subject again for you to-night. While I record my gratitude to Sir James Jeans, I should like to mention also my indebtedness to one of the unsung heroes of astrophysics, Mr F. E. Baxandall, whose work on the spectrum of Beta Lyrae was the beginning of my interest in binaries with unusual spectra. Baxandall, whom I had the pleasure of meeting personally at Cambridge in 1928, was an exceptionally keen observer and we owe to his knowledge and insight the discovery of a system of greatly displaced absorption lines which are visible just before and immediately after the middle of the eclipse of the hot component of β Lyrae by the smaller, cool component. These "satellite" lines are of exceptional importance in many spectroscopic binaries and a large part of my talk to-night will deal with their analogues in other systems.

In 1906 Sir George Darwin, after reviewing his own work on "The Genesis of Double Stars", remarked: "In the sketch which I have endeavoured to give of this fascinating subject I have led my reader to the very confines of our present knowledge. It is not much more than a quarter of a century since this class of observation has claimed the close attention of astronomers; something considerable has been discovered already and there seems scarcely a discernible limit to what will be known in this field a century from now. Some of the results which I have set forth may then be shown to be false, but it seems profoundly improbable that we are being led astray by a Will-of-the-Wisp". He predicted that "probably the observational advance will be more rapid than that of theory" and he concluded that "human life is too short to permit us to watch the leisurely procedure of cosmical evolution, but the celestial museum contains so many exhibits that it may be possible .... to piece together bit by bit the processes through which stars pass in the course of their evolution".

Our astronomical museum has, indeed, grown enormously since Darwin wrote these words, and I shall attempt to give you a brief account of some of the newer exhibits in it.

1. *Interpretation of Previous Results*

Perhaps it is best to start with a well-known system of an eclipsing variable which has been carefully investigated by some of the best photometric workers, including R. S. Dugan at Princeton, and has invariably been found to have a circular orbit. Such a star is U Cephei. It must be remembered that the light-curve yields an accurate determination only of the quantity $e\cos\omega$, which results in a simple manner from the spacings of the total and annular eclipses. In order to determine separately the eccentricity and the longitude of the periastron it is necessary to know also $e\sin\omega$. When the photometric measures are unusually accurate we can determine this quantity from the widths of the two eclipses. For example, when the widths are the same, ω must be approximately 0° and when they are very different, ω is 90° or 270°.

However, a spectrographic study by E. F. Carpenter at the Lick Observatory, in 1930, gave an unsymmetrical velocity-curve from which he deduced $e=0{\cdot}47$ and $\omega=25°$. Hence, $e\cos\omega$ as determined from the velocity-curve would be 0·43, and this, in turn, would mean that the secondary eclipse could not be halfway between two primary eclipses, but should occur about 12 hours later than had been observed with the photometer. Actually, the secondary eclipse cannot be in error by more than a few minutes. Hence, the discrepancy between the light-curve and the velocity-curve was unexplainable in terms of ordinary Keplerian motion.

In 1943 I obtained a new velocity-curve of U Cephei at the McDonald Observatory (**1**). The curve was unsymmetrical and resembled that measured by Carpenter. It is shown in Fig. 1. The ascending branch is not as steep as the descending branch, and the minimum of the curve is broader than the maximum. I have been greatly puzzled by some aspects of this velocity-curve and I have on numerous occasions gone back to it in order to satisfy myself that the interpretation which I had given in 1944 could not be improved upon. It must be noted that the spectral lines of U Cephei are exceptionally poor. The star, whose light we observe at all phases except for a few hours during the total eclipse, is of spectral type B8 or B9. The lines are diffuse and broad because of the rapid rotation of this component around its axis. The enormous velocity of rotation at the equator is shown by the rotational disturbance in the velocity-curve during the partial phases. I was at first perplexed by several points on my curve, just preceding the first contact of the eclipse, because these points all fall systematically below the computed radial velocity of the centre of the B-type star. During the rotational disturbance, this phenomenon is, of course, not observed, but immediately after the last contact the measured points fall systematically too high. Perhaps you will object that the scatter of my measurements is so large that no real significance can be attributed to this result. However, the quality of the spectrum is much better when the star is on the descending branch of the curve than at any other part of its orbit.

I was next puzzled by a strange tendency of my measures to fall systematically above the computed curve between phases 0·8 day and 1·6 day, while immediately after the latter, that is, between 1·6 day and about 1·9 day, there are a surprising number of points below the curve. In 1944 I was under the impression that except for a short interval of time during the partial eclipse of the star, the radial velocities must of necessity refer to the centre of mass of the B-type component and must therefore obey the laws of the two-body motion. However, in the past few years, the picture has been changed, and we have been led to believe that many

close binaries are enveloped in streams of gas which revolve like whirlpools around and inside the system. These streams produce absorption lines when the amount of gas projected in front of the photosphere of the bright binary component is large; they produce emission lines when large masses of gas, excited by the B-type component to emit the Balmer lines of hydrogen, are projected upon the black background of the sky.

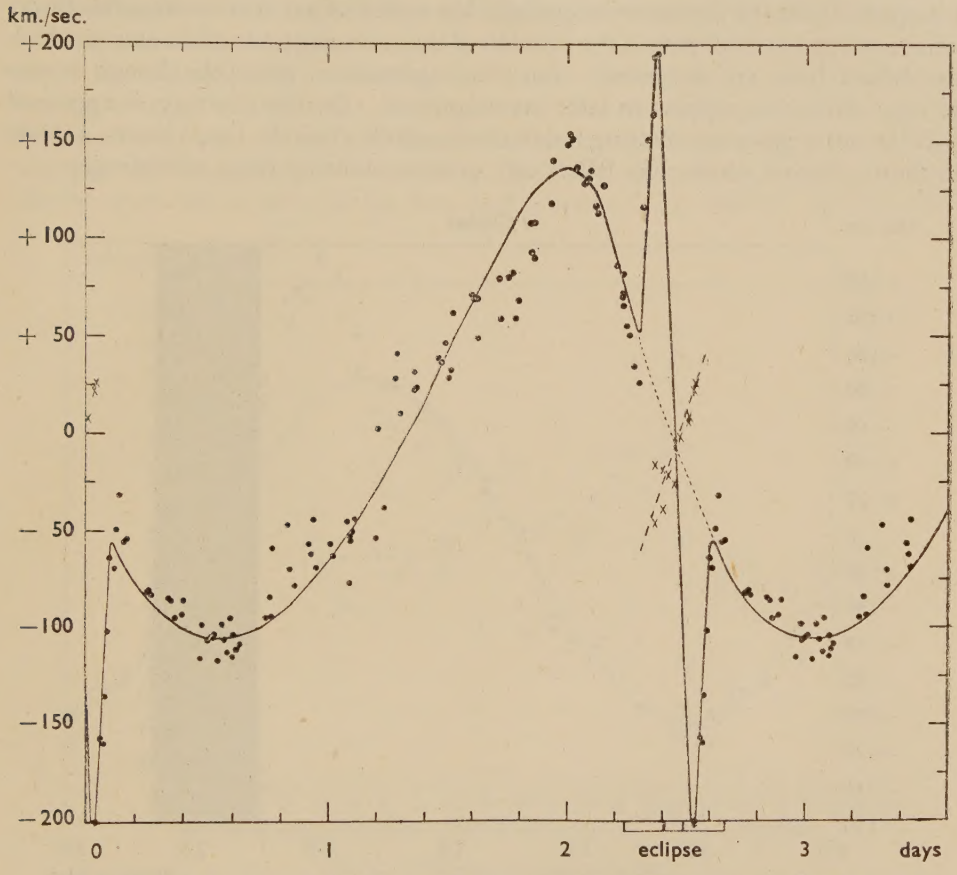

FIG. 1.—*Velocity-curve of U Cephei. The curve is unsymmetrical, corresponding to* $e=0{\cdot}20$, $\omega=40°$. *The photometric observations give* $e\cos\omega=0$.

In the light of these new ideas, it seems reasonable that the velocity-curve of U Cephei is really symmetrical and that it departs from the curve drawn in my article in such a manner as to cross the old gamma axis about 0·1 day earlier. The maximum of the true velocity-curve would then fall approximately at phase 1·83 day, corresponding to a radial velocity of about +80 km./sec. I have tried to sketch this velocity-curve in Fig. 2 and I believe you will notice that the representation of the curve is much better between phases 0·2 day and 1·8 day; but after the latter the observed points are wildly positive. At first this would seem to constitute a serious objection to the proposed hypothesis, but I stated in my earlier investigation that just before first contact and immediately after it the hydrogen lines had strong and fairly sharp violet-displaced cores in absorption, while near the observed maximum of the velocity-curve I suspected that there were cores displaced toward the red side of the broad wings of *H*. It is almost certain now that these displaced cores distort the velocity-curve in the sense of

raising the measured velocities between phases 1·8 and 2·2 days and depressing them after about 2·2 days. This picture is confirmed by the remarkable fact that the hydrogen lines show a much smaller rotational disturbance than do the lines of *He* I, *Ca* II, *Mg* II and *Si* II. The points in Fig. 1 are based upon the latter group of lines and show the maximum amount of disturbance. When I found this peculiar difference between the hydrogen lines and the lines of other elements, I suspected that the former were produced in a shell of gas surrounding the B star whose rotation is smaller than the rotation of the reversing layer of the star in which the fainter lines are produced. But this explanation, plausible though it was in 1944, found no support in later investigations. On the contrary, it appeared that the outer gases are rotating much more rapidly than the inner layers, at least in those systems which, like RW Tauri, possess emitting rings of hydrogen.

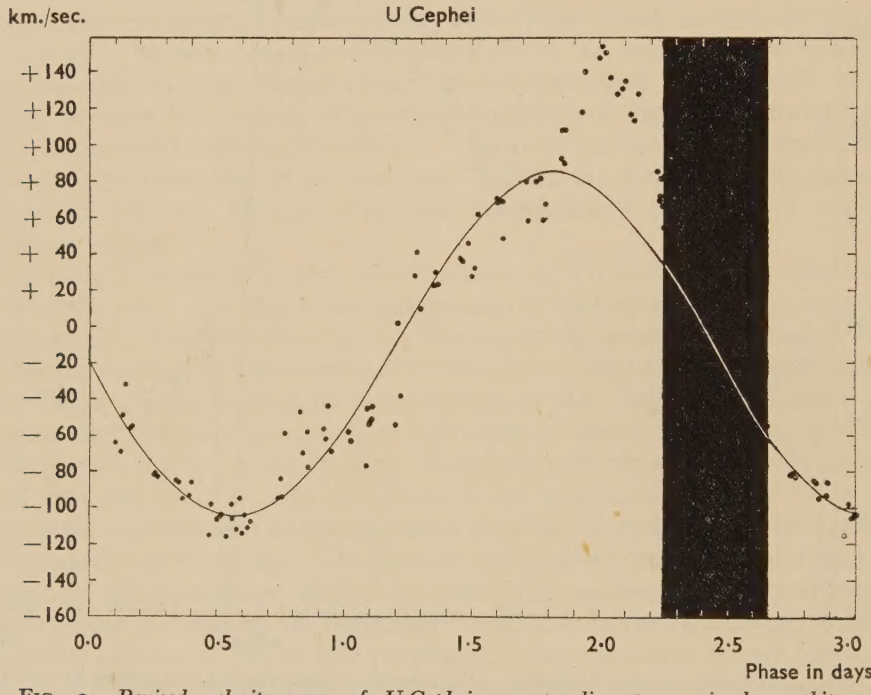

FIG. 2.—*Revised velocity-curve of U Cephei corresponding to a circular orbit with* $\gamma = -15$ *km./sec.;* $K_1 = 95$ *km./sec. The systematic departures from this symmetrical curve are caused by blending with lines produced in a stream of gas surrounding the system. The interval of the eclipse is indicated by a black strip.*

It is now probable that the small disturbance given by the hydrogen lines is caused by a blending of the violet-displaced core which has no rotational disturbance and a stellar hydrogen line produced in the reversing layer in which the rotational disturbance is precisely the same as in the case of *He*, *Ca* II, *Si* II and *Mg* II.

After the end of totality the rotational disturbance is again greatest for the group of fainter lines and is less pronounced in the case of hydrogen, but the difference between the two groups is less striking than before totality. It can be concluded that the unaffected core of the hydrogen lines is now displaced toward the red but is weaker than it was in the first part of the eclipse.

If the case of U Cephei were unique, it might not carry much weight, because admittedly the measures are very uncertain and the conclusions based upon them

must be regarded with caution. However, in 1945 F. J. Neubauer and I made a detailed study of the velocity-curve of the eclipsing variable RZ Scuti. Here the results were even more convincing. The rotational disturbance is enormous, but it is less than half as large when measured in helium as when measured in hydrogen. The rest of the velocity-curve is distinctly peculiar (Fig. 3). The minimum is flat and uniform at about −50 km./sec. The measures indicate a maximum about half a cycle after the minimum, with a velocity in the vicinity of 0 km./sec. As the phases increase, there is at first a tendency for the velocity-curve to decrease, and we suspect that we are dealing with a simple sine curve having a range of the order of 50 km./sec. But the decline in radial velocity does not continue. About two days before the eclipse begins, that is, long before any part of the primary is obscured by the dark secondary, the radial velocity shoots up to about 30 or 40 km./sec. and remains there until the partial eclipse

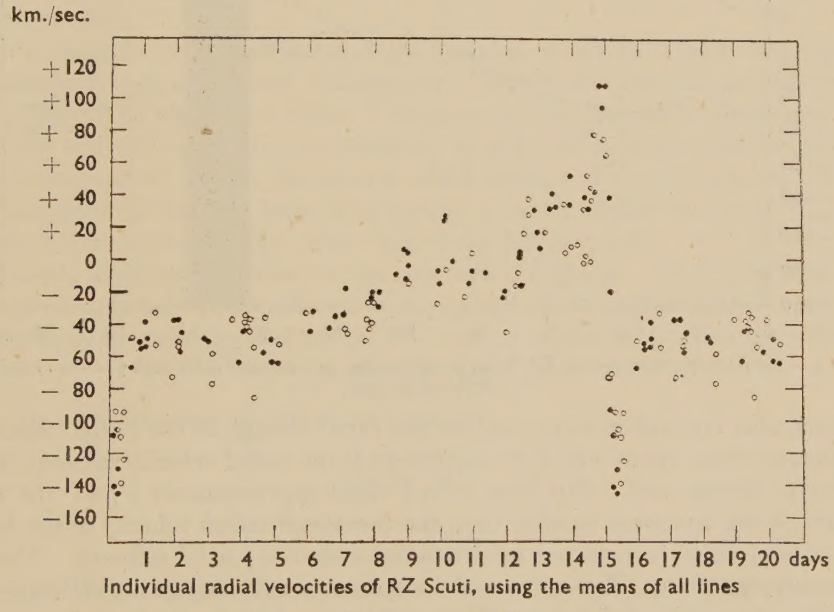

FIG. 3.—*Velocity-curve of RZ Scuti.*

brings with it the normal rotational disturbance of the helium lines. We immediately concluded that in this system stream motion accounts for the blending of spectral lines before the beginning of partial eclipse and that, in this stage, the absorption cores produced by the stream must be displaced toward the red side of the normal stellar absorption line, while in the vicinity of first contact the cores of the absorption lines are displaced toward the violet and produce a marked decrease in the rotational disturbance measured in the hydrogen lines. Fig. 4 shows the McDonald velocities of RZ Scuti, with the true velocity-curve.

Since then, the same phenomenon has been recorded more or less completely in several other systems. Hiltner found it in RY Persei and more recently, together with Hardie, in RW Tauri. I noticed the same succession of events in SX Cassiopeiae and RX Cassiopeiae, and I have suspected it in several other systems, for example, U Coronae Borealis, which I investigated in collaboration with Sahade. Hence, from the published data we can conclude that in all, or nearly all, eclipsing variables whose periods are approximately between 2 days

and 5 days, and in several other systems of longer period, gaseous streams produce blended absorption lines which distort the velocity-curve primarily at phases between 0·8 P and 1·0 P and to a somewhat lesser extent between 1·0 P and 0·2 P. We conclude that either the stream is greater in mass when we observe it prior to the middle of the eclipse or the conditions of excitation and ionization are at that time more favourable to the production of the Balmer lines of hydrogen.

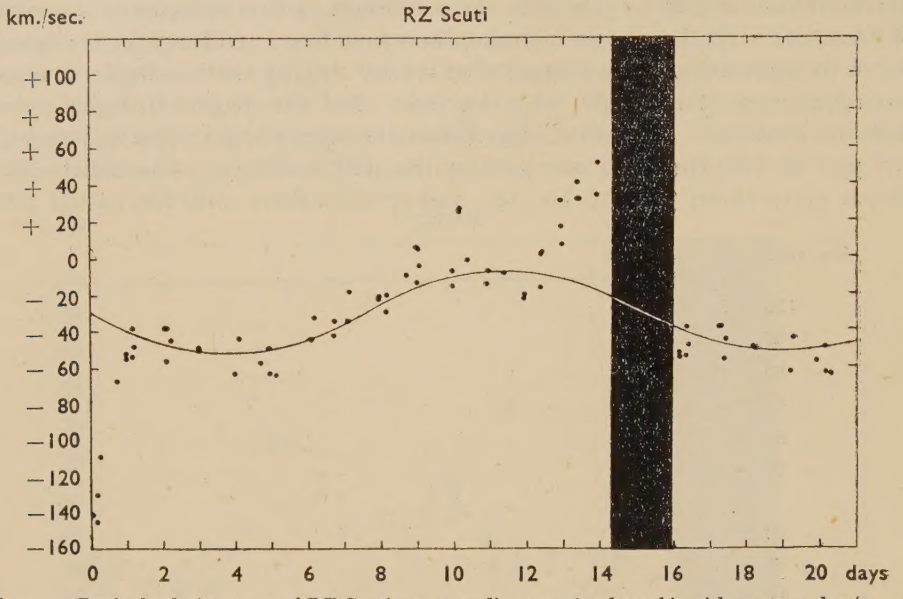

FIG. 4.—*Revised velocity-curve of RZ Scuti corresponding to a circular orbit with* $\gamma = -29$ *km./sec.*; $K_1 = 23$ *km./sec.*

We must also account in some way for the rapid change in the radial velocity of the stream: from about 0·80 P to about 0·85 P the radial velocity exceeds that of the B-type component. But from 0·85 P until approximately 0·95 P the radial velocity of the stream is smaller than the measured radial velocity of the B-type component (which is distorted by rotation during the partial eclipse). This can mean only one thing: The motion of the stream as reflected in the velocity-curve of such a star as U Cephei cannot all be attributed to foreshortening in a uniformly rotating system of coordinates. The observations clearly do not refer to a single mass of gas, because at different phases we observe from different directions, and we therefore record the radial component of the motion, not at the same place in the stream, but each time at that particular portion of the stream which happens to lie in front of the B-type star.

2. *U Sagittae*

The system of U Cephei was the first in which the angular velocity of rotation of the brighter component was found greatly to exceed the angular velocity in the orbit. The same result was later obtained for RZ Scuti and, by Hiltner, for RY Persei. The periods of these systems are 2·5, 15·2 and 6·9 days, respectively. Previously it was believed that in all spectroscopic binaries with $P < 6$ days the periods of rotation and revolution are synchronized, although Swings had already in 1936 called attention to a few exceptions. More recently, Luyten has stated that in the B-type binary α Virginis, P(orbital) = 4 days, the brighter component rotates with P(rotational) = 1·7 day. The observed

value is $V_{\text{rot}} \sin i = 200$ km./sec. If there were synchronism we should have $V_{\text{rot}} \sin i = 85$ km./sec. There is no doubt that the absorption lines are much broader than 2·4 A.—which would correspond to ± 85 km./sec.

Similarly, Luyten, Struve and Morgan found that in AR Cass, $P_{\text{orb}} = 6{\cdot}1$ days, the observed rotation is $V_{\text{rot}} \sin i = 100$ km./sec., while the value calculated for synchronous motions is $V_{\text{rot}} \sin i = 35$ km./sec.

Several of the binaries mentioned have components of very different mass and spectral class. This is not the case in α Virginis, in which the components are of classes B2 and B5, respectively. Despite this exception it is intriguing that several typical Algol binaries, in which the primary component is a massive, hot, luminous and small B or A star while the secondary is a subgiant of small mass, low temperature and large size, show a discrepancy between the angular velocities of rotation of the two components. Thus, in U Cephei the brighter component has $V_{\text{rot}} = 200$ km./sec., while the fainter component has $V_{\text{rot}} \leqslant 50$ km./sec.

With one exception, all typical Algol systems which have been investigated with sufficient care show this discrepancy. While the number is still small, they suggest a cosmogonical cause. Perhaps, in the process of formation of a binary the angular momenta of rotation are distributed in proportion to the mass of each component. Then, in systems which consist of a main-sequence B star and a subgiant G star the latter will exhibit a much smaller value of $V_{\text{rot}} \sin i$, because its size is much larger than that of the hot primary.

The only striking exception to this rule was U Sagittae. A. H. Joy's investigation of its spectrographic orbit in 1930 (**2**) had shown a pronounced rotational disturbance of the velocity-curve of the B8 primary, which led to a determination of

$$V_{\text{rot}}(\text{B star}) = 60 \text{ km./sec.}$$

Since $P = 3{\cdot}38$ days, we find from the relation

$$V_{\text{rot}} = \frac{2\pi r_2}{P},$$

that

$$r_2 = 2{,}790{,}000 \text{ km.}$$

This refers to the small, bright and massive star of the system. The light-curve gives r_1 and r_2 in terms of $a_1 + a_2$, the semi-major axis of the relative orbit:—

	uniform	darkened
r_1	0·292	0·280
r_2	0·175	0·218
$\cos i$	0·095	0·000

Joy found from the radial velocities,

$$a_1 = 3{,}160{,}000 \text{ km.}$$

and

$$m_1/m_2 = 3{\cdot}3.$$

Hence,

$$a_1 + a_2 = 13{,}600{,}000 \text{ km.}$$

and

$$r_2 = 2{,}957{,}000 \text{ km. (darkened).}$$

The agreement between the two methods was entirely satisfactory, and it was therefore believed that

$$P_{\text{orb}} = P_{\text{rot}}.$$

There are no determinations of V_{rot} for the fainter, G-type companion, but Joy remarked that "the lines are remarkably good in view of the large rotation effect which must be present". If we apply to this component the same reasoning as to the primary component, we find from the light-curve that under the assumption of synchronism:

$$V_{rot}(\text{G star}) = 82 \text{ km./sec.}$$

This is entirely consistent with the appearance of the lines observed during the total phase of the eclipse.

Fig. 5 (Plate 10) contains several reproductions from recent spectrograms of U Sagittae obtained at the McDonald Observatory. The original dispersion was 40 A./mm., and the projected slit-width was 0·025 mm. The exposure times range from 4 minutes in full light to about 40 minutes during totality. Experience has shown that rotational broadening of less than 50 km./sec. cannot be detected with this equipment in the observed line-widths. The lines of the G star are very slightly broader than those of the Sun, and we are entitled to conclude that $V_{rot}(G)$ is slightly in excess of 50 km./sec. but definitely less than 100 km./sec. But during full light, phases $0{\cdot}3852\,P$ and $0{\cdot}6249\,P$, the lines of *He* I, *Mg* II, *Ca* II and *Si* II are greatly broadened. This is in agreement with the earlier descriptions of this spectrum. For example, Miss Mary Fowler (**3**) commented on its "broad and very poor lines" and Joy (**2**) stated that "measurement of the spectrum is difficult on account of the width and poor character of the lines". From the lines *He* I 4026, *He* I 3820, *Si* II 3856 and *Si* II 3863 we derive a line-width of about $\Delta\lambda = 3{\cdot}5$ A. or

$$V_{rot}(B) = 135 \text{ km./sec.}$$

This is about twice the value found by Joy. It suggests that either the observed line-contours are broadened by effects other than rotation, or his conclusion must be abandoned that

$$P_{orb} = P_{rot}(B).$$

In order to investigate this question I have obtained an extensive series of spectrograms of U Sagittae, especially during the first half of the eclipse on 1949 August 2. The results are given in Tables I and II and in Figs. 7, 8, 9 and 10. The heliocentric phases were computed with the help of Jacchia's formula (**4**):

$$\text{Principal Minimum} = \text{JD } 2{,}417{,}130{\cdot}4151 + 3{\cdot}3806184\,E + 0{\cdot}0043 \sin(0°{\cdot}107\,E + 266°).$$

The observations show that on August 2 the middle of the principal eclipse occurred about 45 minutes earlier than would be consistent with Jacchia's elements. On spectrograms CQ 7764 and 7772 the G star is about equally faint.

TABLE I

Star-lines measured in U Sagittae

H	*H*	*He* I	*Si* II	*Mg* II
3697·15	3770·63	3819·62	3862·59	4481·23
3703·86	3797·90	4026·20	4128·05	*Ca* II
3711·97	3835·39	4471·50	4130·88	3933·67
3721·94	3889·05	*Si* II	*Fe* II	...
3734·37	4101·74	3853·66	4233·17	...
3750·15	4340·47	3856·02	...	...

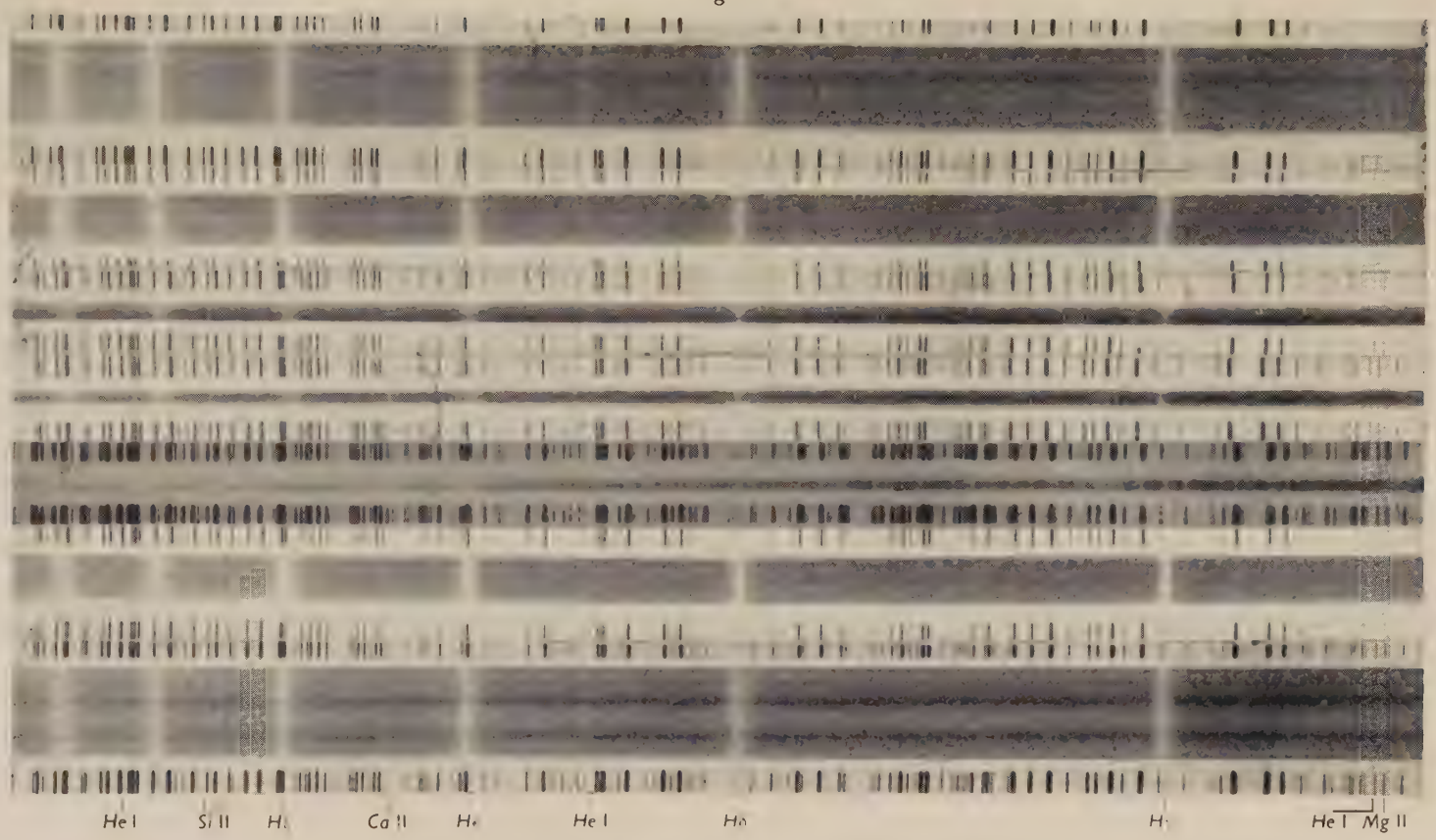

FIG. 5.—*The spectrum of U Sagittae showing broad and diffuse absorption lines of He I, Mg II, Si II and Ca II outside of eclipse. Notice double absorption lines of H at phases 0·9660 P and 0·9737 P.*

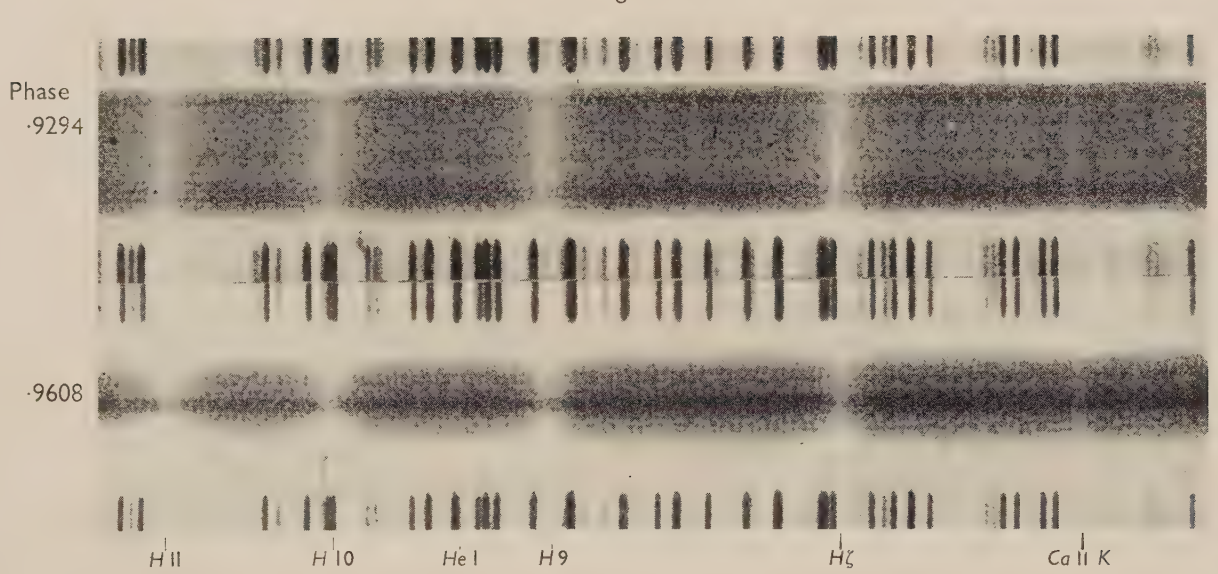

Fig. 6.—*Two spectrograms of U Sagittae showing sharp absorption cores displaced toward the violet side of the centre of the broad wings.*

Otto Struve, George Darwin Lecture.

TABLE II

Radial Velocities of U Sagittae

Plate CQ	Date (1949) U.T.	Exposure time (minutes)	Helio-centric phase	Hydrogen I	Hydrogen II	*He*	All lines except *H* and *Ca* II	*Ca* II I	*Ca* II II
7687	July 30·1083	20	0·0318	−48·7	...	− 69·5	− 61·9	−36·5	...
88	30·1226	19	0·0360	−48·7	...	− 54·0	− 58·1	−48·7	...
89	30·1344	15	0·0395	−56·2	...	− 57·3	− 56·0	−45·7	...
90	30·1452	15	0·0430	−53·6	...	− 75·2	− 63·7	−42·6	...
91	30·1559	12	0·0459	−52·4	...	− 55·8	− 44·0	−36·6	...
92	30·1653	16	0·0487	−42·1	...	− 39·3	− 44·9	−45·8	...
93	30·1750	10	0·0515	−46·6	...	− 71·1	− 58·6	−58·0	+131·8
94	30·1830	9	0·0539	−44·9	...	− 44·6	− 54·4	−51·9	...
95	30·1889	10	0·0557	−47·7	...	− 58·2	− 48·1	−30·6	...
96	30·1972	10	0·0581	−49·4	...	− 39·4	− 37·9	−36·7	...
99	30·2597	8	0·0766	−52·9	...	− 60·9	− 48·9	−39·9	...
7700	30·2648	5	0·0781	−44·8	...	− 42·4	− 38·8	−33·8	...
01	30·2690	5	0·0794	−44·6	...	− 40·2	− 32·5	−43·0	...
05	30·3760	15	0·1110	−59·9	...	− 51·4	− 48·1	−37·8	...
06	30·3868	10	0·1142	−61·0	...	− 71·9	− 74·2	−43·2	...
12	31·1938	6	0·3529	−70·4	...	− 58·6	− 70·5	−61·5	...
13	31·1986	6	0·3543	−65·0	...	− 71·9	− 66·9	−34·0	...
14	31·2035	6	0·3558	−62·4	...	− 63·3	− 62·4	−49·3	...
17	31·3031	5	0·3852	−60·9	...	− 37·9	− 44·5	−53·4	...
18	31·3090	8	0·3870	−56·4	...	− 53·9	− 65·1	−65·6	...
19	31·3139	5	0·3884	−52·2	...	− 65·7	− 58·9	−41·2	...
26	Aug. 1·1031	19	0·6219	+35·9	...	+ 40·5	+ 46·6	+29·0	...
27	1·1132	8	0·6249	+44·5	...	+ 21·0	+ 24·4	+ 4·7	...
28	1·1354	14	0·6314	+41·3	...	+ 15·7	+ 18·7	+10·8	...
33	1·3944	6	0·7081	(+79·8)	...	+ 40·7	+ 49·0	+ 4·1	...
38	2·1028	8	0·9176	+40·2	...	+ 5·4	+ 26·2	+11·2	...
39	2·1083	6	0·9192	+34·5	...	+ 23·9	+ 32·4	+ 2·0	...
40	2·1125	6	0·9205	+28·4	...	+ 13·5	+ 22·1	+ 8·1	...
41	2·1171	5	0·9218	+41·2	...	+ 13·1	+ 35·7	+14·2	...
42	2·1219	7	0·9232	+27·9	...	+ 22·7	+ 45·3	+ 8·1	+191·7
43	2·1271	6	0·9248	+31·0	...	+ 44·4	+ 42·3	+ 2·0	...
44	2·1316	5	0·9261	+36·2	...	+ 9·1	+ 19·8	− 4·1	...
45	2·1427	5	0·9294	+23·8	...	+ 22·9	+ 33·0	+14·1	...
46	2·1465	6	0·9305	+26·4	...	+ 13·2	+ 20·2	−10·3	...
47	2·1531	6	0·9325	+34·0	...	+ 18·5	+ 34·2	+23·2	...
48	2·1601	7	0·9345	+20·8	...	+ 4·5	+ 20·4	+ 1·9	+219·0
49	2·1660	8	0·9363	+28·9	...	+ 28·0	+ 20·8	+ 5·0	...
50	2·1708	6	0·9377	+26·4	...	+ 31·7	+ 32·8	+ 1·9	...
51	2·1750	6	0·9390	+30·0	...	+ 46·8	+ 36·7	+20·2	...
52	2·1861	8	0·9422	+27·6	...	+ 31·0	+ 40·2	+23·2	...
53	2·1934	9	0·9444	+27·9	...	+ 28·0	+ 32·6	+23·2	+253·2
54	2·1997	11	0·9463	+32·1	...	+ 53·9	+ 50·3	+23·1	+191·5
55	2·2069	8	0·9484	+34·6	...	+ 18·5	+ 38·9	+17·1	...
56	2·2139	10	0·9505	+29·2	...	+ 54·6	+ 40·8	+17·1	+188·4
57	2·2229	14	0·9531	+32·5	...	+ 38·7	+ 41·3	+11·0	...
58	2·2333	14	0·9562	+38·5	...	+ 42·5	+ 49·9	+23·1	...
59	2·2417	8	0·9587	+28·0	...	+ 46·5	+ 54·3	+39·0	...
60	2·2490	13	0·9608	+29·3	...	+ 39·8	+ 43·2	+39·0	...
61	2·2583	10	0·9636	+28·8	...	+ 60·4	+ 62·6	+33·4	+174·3
62	2·2663	13	0·9660	+ 6·9	+212·1	+ 53·1	+ 51·7	+29·1	+231·7
63	2·2747	9	0·9684	+10·5	+199·1	+ 53·2	+ 55·2	+49·4	+226·8
64	2·2819	12	0·9706	+13·1	+220·4	+ 62·4	+ 67·1	+24·2	+229·8
65	2·2924	16	0·9737	+13·8	+242·9	+ 62·5	+ 68·2	+18·1	+248·0
66	2·3028	18	0·9768	+16·0	+255·3	+ 67·3	+ 67·3	+43·2	+251·1
7772	2·4292	30	0·0142	−51·8	...	−164·2	−154·5	−76·8	...

The corresponding phases are 0·9706 *P* and 0·0130 *P*. Hence, the shift is 0·0082 *P* = 0·028 day = 40 minutes. But a larger shift results from the radial velocities of the *H* lines in Fig. 7, at phases 0·625 *P* and 0·375 *P*. The curve is taken from Joy's paper. Agreement would be obtained if this curve were shifted

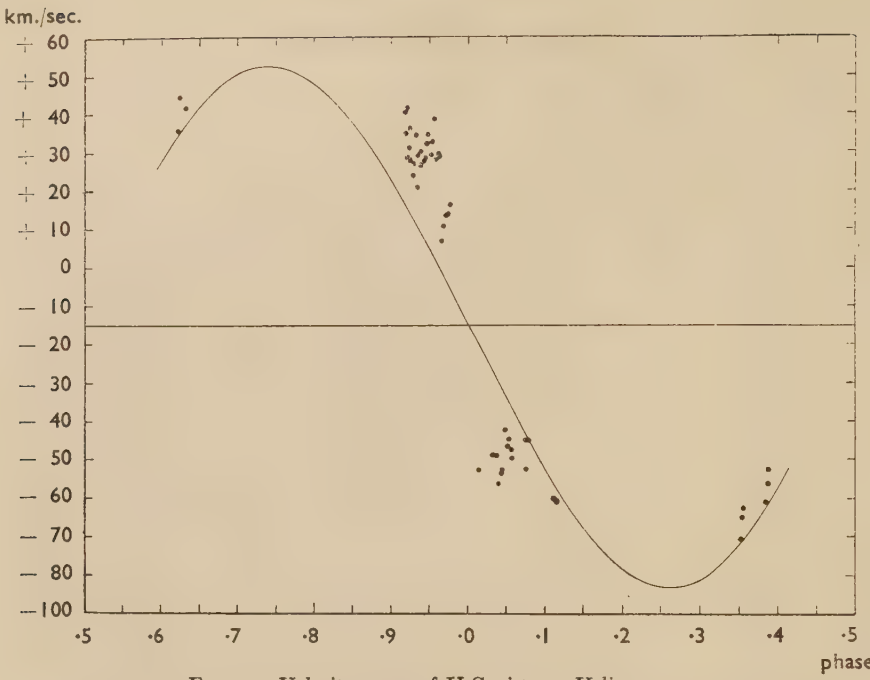

FIG. 7.—*Velocity-curve of U Sagittae: H lines.*

The curve is from Joy's orbit. The middle of eclipse on 1949 August 2 occurred 0·028 day earlier than was predicted by the elements.

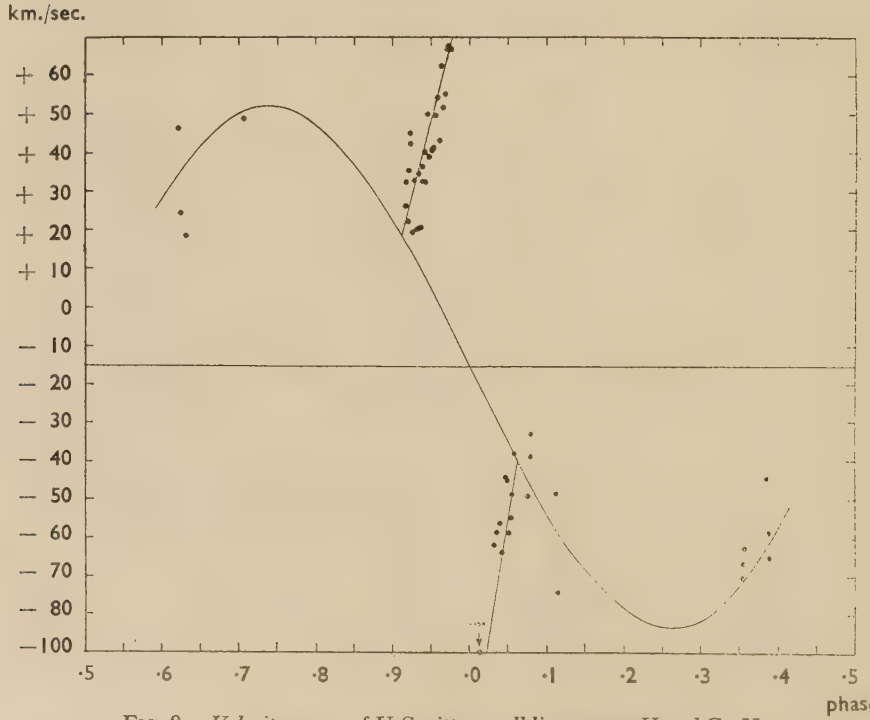

FIG. 8.—*Velocity-curve of U Sagittae: all lines except H and Ca II.*

The curve is from Joy's orbit. The middle of eclipse on 1949 August 2 occurred 0·028 day earlier than was predicted by the elements.

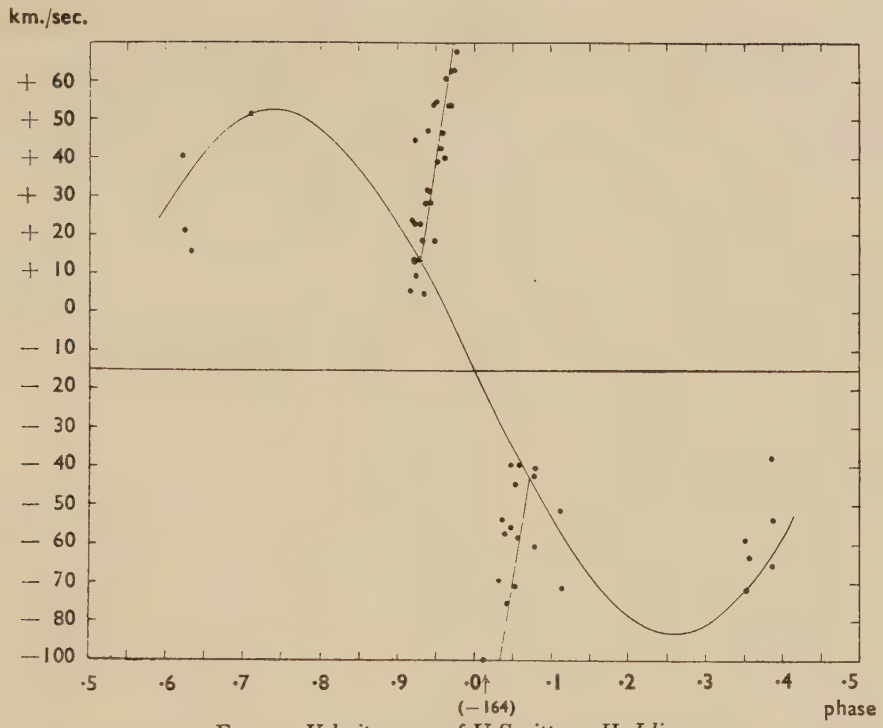

FIG. 9.—*Velocity-curve of U Sagittae: He I lines.*

The curve is from Joy's orbit. The middle of eclipse on 1949 *August* 2 *occurred* 0·028 *day earlier than was predicted by the elements.*

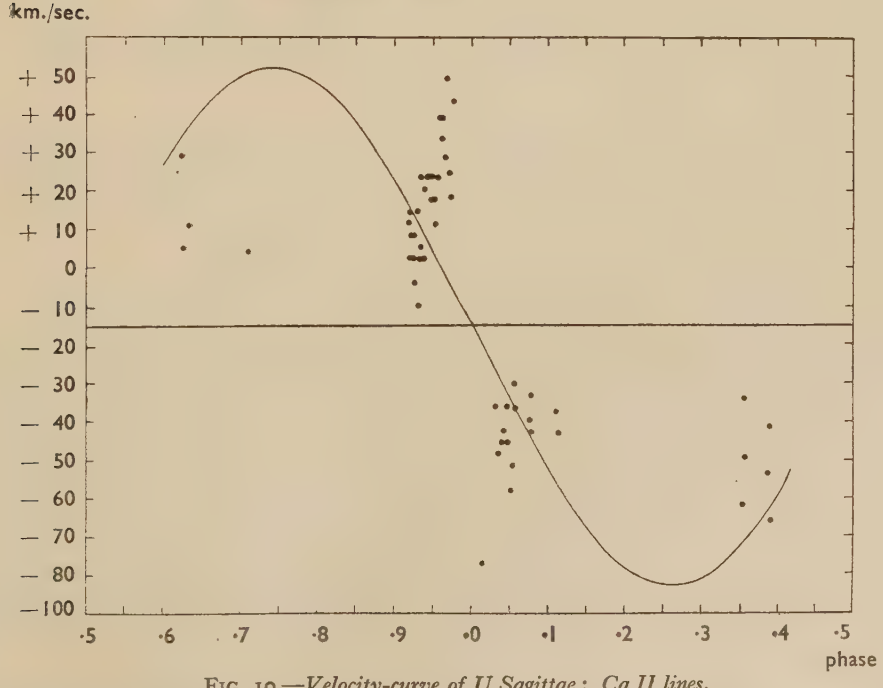

FIG. 10.—*Velocity-curve of U Sagittae: Ca II lines.*

The curve is from Joy's orbit. The middle of eclipse on 1949 *August* 2 *occurred* 0·028 *day earlier than was predicted by the elements.*

to the left by about 0·015 *P*. The velocities during the partial phases of the eclipse cannot be used for this comparison.

After these estimates had been made, Dr N. L. Pierce of Princeton informed me that from a compilation of various photometric data he believes that the period of U Sagittae has become much shorter and may, in fact, be given by

$$\text{Principal Minimum} = \text{JD } 2{,}425{,}125{\cdot}5798 + 3{\cdot}3806122\,\text{E}.$$

The most recent observation available to him was a visual estimate by Szafraniec in 1947 which gave $\Delta = -0{\cdot}020$ day at about $\text{E} = 4500$. My spectrographic observations, at $\text{E} = 4732$ and $\Delta = -0{\cdot}028$ day, indicate that the period may now be even shorter. Dr J. Irwin, at the Goethe Link Observatory, obtained on October 2·0 and 9·0 a correction of −0·0162 day to the linear formula:

$$\text{Hel. Min.} = 2{,}417{,}130{\cdot}4170 + 3{\cdot}3806184\,\text{E},$$

from two series of photoelectric observations.

It is certain that Jacchia's sine curve must be abandoned, and with it also the conclusions of Tchudovitchev (**5**) regarding a 30-year period of the rotation of the line of apsides. The spectroscopic observations are not accurate enough to give a reliable determination of e, and I agree with Luyten that the best we can now do is to assume that the orbit is circular. The secondary eclipse of U Sagittae is very shallow and the photometric observations cannot be used to give $e \cos \omega$.

Perhaps it is most instructive to consider first Fig. 8, representing all lines except *H* and *Ca* II. The rotational distortion is very conspicuous and is much larger than that given by Joy. The difference results probably from the fact that he did not distinguish between the different elements. Extrapolating the measures to the beginning of totality we find a maximum rotational disturbance of the order of 75 to 80 km./sec. at the second contact. The corresponding value for the third contact is very uncertain, but we may estimate it at 135 km./sec. Allowing for the uncertainty of the latter value, we estimate from the disturbance a maximum effect of about 85 or 90 km./sec.

But even this value is undoubtedly somewhat underestimated. It has already been shown by R. O. Redman (**6**) that the *H* lines of U Sagittae and U Cephei become weaker at the limbs of the two B-type stars, and for the latter this was confirmed at McDonald, not only for *H* but also for *He* I and *Mg* II. I am now able to make the same confirmation in the case of U Sagittae.

As the partial eclipse progresses, the lines of *He* I, *Mg* II and *Si* II at first become noticeably narrower and deeper. Their contours resemble those shown on the left side of Fig. 2 of *Ap. J.*, **99**, 234, 1944, where a large dark star is assumed to eclipse a small bright star. Since the equivalent width of each line remains the same, the contour must deepen as it becomes narrower. This process continues until about phase 0·97 *P*, or until about $-0{\cdot}02\,P = 1^{\text{h}}\,35^{\text{m}}$ before second contact. At this time, we see in Fig. 11 that only slightly more than one-half the B star is covered by the G star. Yet, from 0·97 *P* to 0·99 *P*, where totality sets in, the lines of *He* I, *Si* II and *Mg* II not only fail to increase in sharpness but they actually become more difficult to see: their equivalent widths become smaller and their widths increase. This has no relation to the light of the G star, whose continuous spectrum is completely negligible until about 30 minutes before and after totality.

We must conclude that these lines have very weak and indistinct contours except near the centre of the B star's disk. Whether this is caused by "true

absorption", as was suggested by Redman, or by something else, such as electron scattering, cannot be decided. The observed weakening of the lines toward the limb shows that we never really observe the full amount of the rotational disturbance, but only a mean value, heavily weighted in favour of the central regions of the star's disk. It is impossible to allow for this effect quantitatively, but perhaps we shall not be far off the mark if we adopt

$$V_{\text{rot}}(\text{B}) = 100 \text{ km./sec.}$$

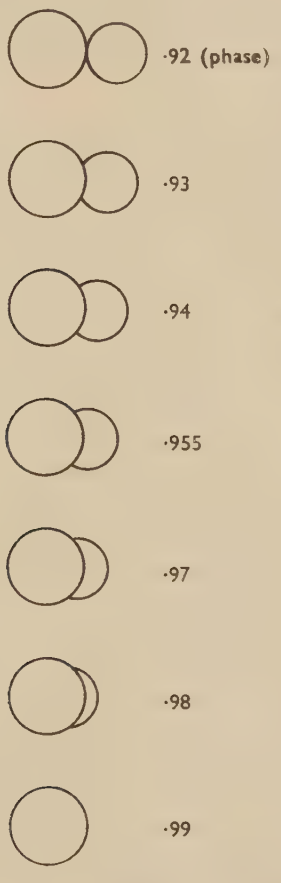

FIG. 11.—*Relative position of the two components of U Sagittae during the eclipse.*

Let us inspect the velocities given by the *H* lines. Since my plates extend far into the ultra-violet I was able to measure 12 members of the Balmer series. Despite the great widths of these lines, the measures are reasonably accordant and the scatter is much less than in previous work on this star. Between phases 0·9176 *P* and 0·9636 *P* we find a compact group in Fig. 7 which lies about 20 km./sec. above Joy's velocity-curve. Between these same limits of phase the velocities from *He* I alone, or from *He* I, *Si* II, *Mg* II and *Fe* II combined, show a group of points, in Figs. 8 and 9, which on the average is not more than 10 km./sec. above the curve. More surprising still is the absence of any tendency of the *H* lines to follow the pronounced straight-line trend of the rotational disturbance seen in Fig. 8 and Fig. 9.

During this interval in phase the *H* absorption lines are conspicuously unsymmetrical. They possess deep and fairly sharp cores which are displaced

towards the violet side of the centre of the broad wings (Fig. 6, Plate 11). This is precisely the same effect as that previously found in U Cephei and several other stars. But if the core is already displaced 20 km./sec. towards the red from Joy's velocity-curve, then we must infer that the *H* gases giving rise to the wings must be displaced even more; 150 km./sec. would not be too much, since a smaller displacement would not be so readily seen, even without measurement. But the rotational disturbance is only of the order of 10 km./sec. at these same phases. Thus, the wings of *H* are produced in a layer that cannot be identical in its properties with the rotating layer in which the lines of *He* I, *Si* II and *Mg* II are produced. This is an interesting puzzle which I shall try to solve after my return to the United States.

After phase 0·9636 *P* we have five spectrograms on which the *H* lines are distinctly double. This is shown in Fig. 5. On the average, the red component is the weaker and the broader of the two, and its velocity is +220 km./sec. The violet component is identical with the violet-displaced core observed prior to phase 0·9660 *P*. Apparently it follows the general trend of the velocity-curve and is not subject to a rotational disturbance. It persists to the very beginning of totality. The red component is perhaps identical with the line which we had attributed to the wings before phase 0·9660 *P*. It is displaced too much to be caused by any part of the rotating B star. Both components which we have measured in the *H* lines have their origin in a stream of gas.

The normal stellar *H* lines from the B star are probably predominant in the spectrum of U Sagittae at all phases between 0·2 *P* and 0·8 *P*. Joy's velocity-curve and our measures at a few phases suggest no detectable cores in the *H* lines and there is no appreciable distortion of the curve, such as has been observed in U Cephei and RZ Scuti: either the stream is not sufficiently dense to show these cores, or their velocities coincide with those of the B star.

The usual lines of *He* I and *Mg* II are not affected by the stream, and there is no indication of the wide splitting observed in *H*. The ultra-violet lines of *Si* II may show the splitting, but the evidence is not conclusive. The line *Ca* II K shows the splitting conspicuously—even more so than do the *H* lines. On the descending branch of the velocity-curve the *Ca* II line is much stronger and sharper than elsewhere. Between first and second contact it is often double, and the red component gives the same velocity as does the red component of *H*. It is present on most of my plates of August 2, and it confirms our previous contention that the *H* lines also possessed such a red-displaced component a short time prior to phase 0·9660 *P*. Only in the case of *H* do the normal stellar lines of the B star obscure the structure of the pair, probably because of their Stark broadening.

The violet component of *Ca* II shows the rotational disturbance, but all measures of *Ca* II (Fig. 10) indicate a tendency to lie closer to the gamma-velocity than do the other lines. This must be due to blending with an interstellar *Ca* II line. These measures are, therefore, not reliable for a discussion of stream motion in the system.

The blend *H* + *He* I at λ 3889 shows strikingly the components of the lines originating in the stream. Thus, the unsymmetrical core with its violet displacement is better visible in this line than in any other member of the Balmer series. This is undoubtedly caused by the *He* I line, whose lower level is metastable and which is therefore very useful in detecting dilution of radiation. The stream in U Sagittae shows some effect of dilution, but it is not possible to measure it accurately.

The observations between third and last contacts are not numerous enough to draw many conclusions. But we may indicate a few: the cores of *H* are strong and are approximately centrally placed over the wings. This is similar to U Cephei but different from RW Tauri. No doubling of the *H* lines has been observed. The stream of large displacement is not visible.

Fig. 12 represents the system of U Sagittae as seen by an observer who revolves around it in a direction opposite to that of the orbital motion. The partial eclipse begins at 0·92 *P* and ends at 0·08 *P*. Totality begins at 0·99 *P* and ends at 0·01 *P*.

It is not possible to indicate the orbit of the stream of gas, but Kuiper has given some theoretical reasons to believe that it may flow around the system in the direction of the orbital motion.

At this stage we cannot compute the actual orbits of the atoms. For one thing, the initial coordinates and velocities are not known; for another, we are not even certain that gravitational forces alone are active. Perhaps there are, in addition, electrical and magnetic forces. Hence, we shall proceed in a purely observational manner and propose a model which meets the observations without attempting to justify it theoretically.

Let us suppose that the flow occurs mostly between two successive equipotential curves in the *XY*-plane, given by

$$\frac{1-\mu}{r_1}+\frac{\mu}{r_2}+\frac{x^2+y^2}{2}=\frac{C}{2},$$

where $\mu=\dfrac{m_2}{m_1+m_2}$ and r_1, r_2 and C have their usual meaning (**7**).

Let us assume that the motion is uniform, or approximately uniform, between the two inner curves in Fig. 12. Then at all phases between about 0·15 *P* and 0·85 *P* the line of sight would project upon the disk of the small B star different sections of a formation which is very nearly circular in the *XY*-plane and, therefore, shows no radial component of velocity between the stream and the centre of the B star. Only if the gases were excited to emit *H* lines could we observe double bright lines, as we actually do in RW Tauri, SX Cassiopeiae, RX Cassiopeiae, RZ Ophiuchi, etc. The velocities corresponding to these motions are of the order of 200 km./sec., but they are not the same for all stars, varying about as $V^3 \propto P^{-1}$. This is sufficient to enable us to identify the large displacement of the stream lines in U Sagittae with the ring-like formation studied by Joy in RW Tauri and by me in many other systems. The only reason we fail to see bright, double *H* lines in U Sagittae and U Cephei is that the system is too small: in a larger system the dimensions of the stream are increased, but the radii of the stars are not larger in the same proportion.

At phases between 0·97 *P* and 0·99 *P* the line of sight cuts through two lobes of the stream: that part of it which hugs the advancing hemisphere of the G star has a large velocity of recession; presumably it gives rise to the components with the velocity of +200 km./sec. The other part, near the B star, also shows recession, but the direction is almost at right angles to the line of sight and, accordingly, the resulting line should lie only slightly above the velocity-curve defined by the lines of *He* I, *Mg* II and *Si* II which come from the reversing layer of the B star.

After the end of totality the phenomena should repeat themselves, with all displacements having the opposite signs. Actually, our observations show less pronounced stream lines. This may indicate temporary or permanent differences in the properties of the two sides of the stream. But it is of interest that in my

1943 observations of U Cephei there were double lines of *H* and *Ca* II after totality, and the entire picture could be very nearly duplicated with the stream stronger after mid-eclipse than before it.

Different stars show distinguishable differences in the properties of the streams. This may be caused in part by differences in $\mathfrak{m}_1/\mathfrak{m}_2$. For example, in U Cephei it is not as extreme as in U Sagittae, and the two halves of the equipotential curves would be more nearly alike. But in both systems the fainter, large component nearly fills the space provided by the first critical equipotential surface in Fig. 12.

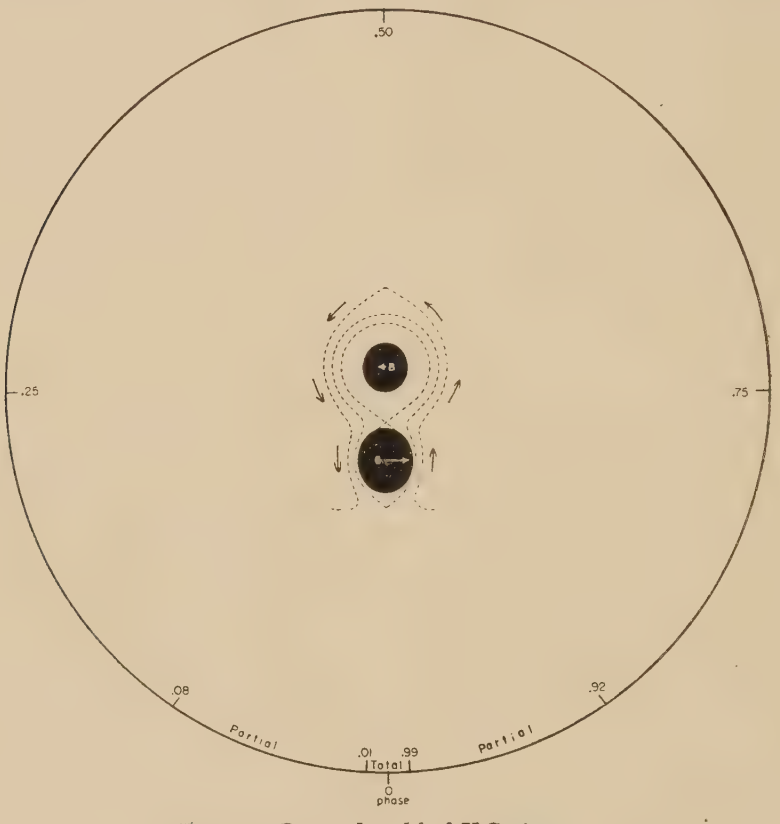

FIG. 12.—*Proposed model of U Sagittae.*

The streams of U Sagittae bear a resemblance to the inner system of streams in β Lyrae and the stream lines are identical with the "satellite lines" of Baxandall's work.

As we have already indicated, the proposed model of a stream does justice to the observations in U Sagittae and in various other stars, but it is difficult to explain theoretically. If gravitation (and radiation pressure) alone is considered, the large, approximately circular space around the B star will probably be filled with matter, and the inner regions of this gas will produce periodic stream-motion in a ring around B. This is essentially Kuiper's ejection mechanism of type A. The stream from the G star will collide with the ring, and will be deflected to produce an outer layer somewhat as we have pictured.

Incidentally, the ring-theory of the gaseous stream around the B star adds weight to W. Lohmann's attempt to determine the mass of the central star in such systems as RZ Ophiuchi and SX Cassiopeiae (**8**).

The observations of the rotating rings are difficult, and for any single star they are often quite incomplete. But we can illustrate the problem we are trying to solve by means of a schematic diagram representing a composite velocity-curve derived from a variety of physically similar systems. Fig. 13 shows such a diagram. The velocities are referred to the centre of the hot component (the B star in U Sagittae). Let us consider only that part of the stream which we can observe between phases $0{\cdot}5\,P$ and $1{\cdot}0\,P$. From the double emission lines (RW Tauri, RS Cephei, etc.) we infer that the real velocity of the ring is roughly 200 km./sec. and is at least approximately constant. Now start with phase $0{\cdot}5\,P$. At this point the radial velocity is zero. We do not know the location of the absorbing gas and we must make a guess, based upon the eclipses of the bright components and the information that dilution, though present, is not large.

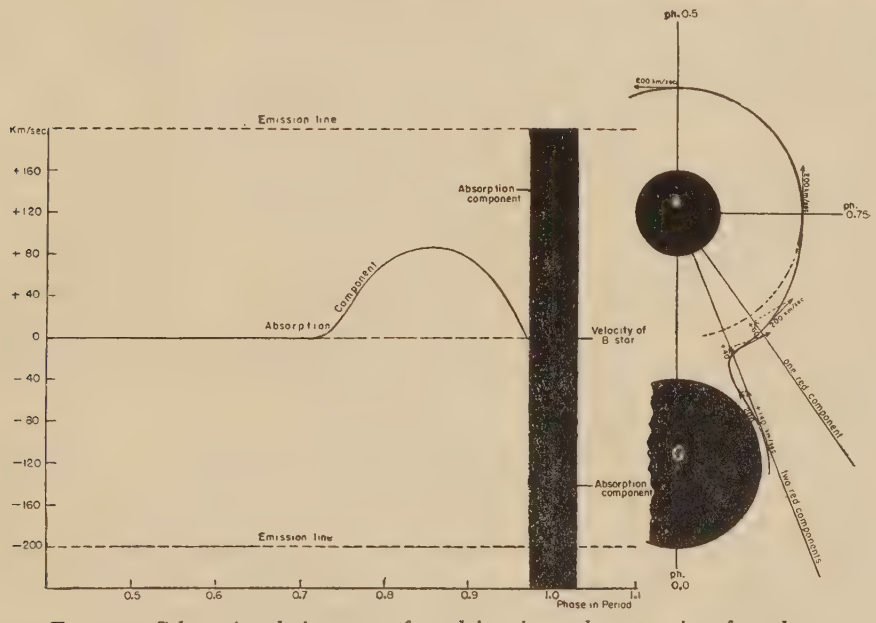

FIG. 13.—*Schematic velocity-curve of revolving ring and construction of envelope.*

Take $r(0{\cdot}5\,P) = 2\,R(\mathrm{B})$. Then, until about phase $0{\cdot}75\,P$ there is no change in V, and hence r remains the same: the formation, up to this point, is circular in cross-section. From here on V is positive, which means that the distance r from the B star increases. If we were observing a fixed portion of the ring, we could simply integrate the velocity curve and derive r. But we do not observe at each phase the same mass of gas. If we fix our attention on a certain condensation of the gas and follow it along its path, we notice that it covers the angle between $0{\cdot}75\,P$ and $0{\cdot}99\,P$ approximately in the interval

$$t_{\mathrm{sec}} = \frac{2\pi r}{200} \times 0{\cdot}24\,P.$$

The B star may have a radius

$$r = 2 \times 10^6\ \mathrm{km}.$$

Hence

$$t = 1{\cdot}4 \times 10^4\ \mathrm{sec}.$$

During this interval the condensation changes its distance from B (if we pass from $0{\cdot}98\,P$ to $0{\cdot}75\,P$) at first slowly, then more rapidly, and finally again quite slowly. A rough integration yields a change in the distance of the order of 10^6 km., or $\frac{1}{2}r$.

But near phase 0·95 P we observe not only one absorption line, but also another whose radial velocity is ~140 km./sec. We have no knowledge of the location of the gas producing this component, but from the assumption that the flow is uniform at 200 km./sec. we infer that we observe this mass almost in the direction of flow. It is not possible to locate such a stream inside the ring-like feature we have been discussing. Hence, it must be located outside, and the only plausible location is along the advancing limb of the G-type component. Thus, we return essentially to the picture of Fig. 12.

3. *Origin and Evolution of Close Binaries*

The ring-like structures which we have described appear to form a general property of all close binaries, including those of the W Ursae Majoris type. We shall designate the resulting model of a binary in a common "whirlpool" of gas as the "peanut model" (following a suggestion by V. Kourganoff), but we must remember that the "shell" of this peanut is present only in the plane of the orbit. Its extent in the Z coordinate is not known, but is probably small.

There are indications in several systems that the stream is unstable and is being dissipated into space. In β Lyrae and HD 47129 (J. S. Plaskett's star of large mass) this dissipation can be measured as a general expansion of the gases constituting the streams. In one system at least, namely UX Monocerotis, we can observe the light of the G-type star at all phases. There are indications of violent motions away from the star, principally in the direction of the orbital motion but also in the direction away from the A-type star. These motions are subject to rapid changes, as are also the intensities of the corresponding lines. Perhaps this is prominence activity on a very large scale, which, for some obscure reason, predominates on the advancing sides of the two stars.

Long ago S. I. Bailey of the Harvard Observatory discovered a remarkable periodic change in the line intensities of the B-type binary μ Scorpii. The lines of the approaching star were relatively stronger than the lines of the receding star. This phenomenon has since been found in a large number of close systems. It is conspicuous in Plaskett's star (HD 47129), AO Cassiopeiae, α Virginis, σ Aquilae, β Scorpii—all of very early spectral class and of short period, and it is also observed in all W Ursae Majoris systems thus far investigated. It can be observed best in systems having double lines in their spectra, and the variation affects mostly the fainter component. But there are indications that the stronger components are also affected.

The simplest explanation of this phenomenon is in terms of the thickness of the absorbing mass of gas in the streams. This would mean that the envelope produced by the stream is thicker over the advancing side of each component. The shell of our peanut looks distorted, as though it were rotated slightly in the forward direction against the line joining the centres of the two stars.

My principal conclusion is that all close binaries possess these common envelopes. Since the envelopes cannot be dynamically stable they probably play a role in the evolution of the binaries. It is therefore desirable to summarize the observational information which has been obtained during the past few years:—

1. The envelopes form ring-like structures flowing around the system with velocities of several hundred km./sec. These rings are located in the orbital planes. From the dilution effect we estimate that their extent is not larger than two or three times the diameter of the brighter star, and from the eclipses of the brighter components in the spectrum we obtain a similar scale of dimensions.

2. From the variation in the intensities of double absorption lines in close systems we infer that the thickness of the ring is greater in the direction of motion of each star.

3. In several systems expansion of the ring has been observed, but in the majority it is not measurable. A range in the velocity of expansion between 0 and 100 km./sec. is indicated. Observational limitations would prevent us from observing the expansion when it is less than about 10 km./sec.

4. In several systems there are erratic variations in the velocities of the gases in the shell, suggesting that they are perhaps related to prominences. There are also erratic variations in the radial velocities of the lines produced in the shell, strengthening still further the hypothesis of prominences on the forward sides of the two stars.

5. In the system UX Monocerotis violent motions away from the G-type subgiant component are considered as direct evidence of prominence activity. These phenomena are also observed in several Wolf-Rayet binaries. Explosive motions are thus frequent and in the case of UX Monocerotis they seem to be concentrated on one hemisphere, producing a gradual diminution in the mass of this star and probably also in its angular momentum.

6. In the case of Beta Lyrae the G-type component is small in size. This permits us to observe the absorption effects of the stream with unusual clarity. The results are the satellite lines discovered by Baxandall. There is also a general expansion of the ring of Beta Lyrae.

7. In systems with deep eclipses and relatively large dimensions double emission lines are the rule. Their eclipses were first observed by A. H. Joy in RW Tauri. Variations in the relative intensities of the bright components and even in the total intensity of the emission features show that the phenomena are irregular in character and that the flow does not have the continuity of a permanent feature, but resembles an effect connected with prominence activity. The rotating rings of Joy are believed to be identical with the rings observed in absorption.

8. There is a general trend among spectroscopic binaries having periods between two and five days to give distorted velocity-curves simulating large eccentricities and values of ω which are in the vicinity of 40°. The statistical discussions by Miss Scott and others have shown that a large number of spectroscopic eccentricities must be spurious.

9. In a few systems the distortion of the velocity-curve can be directly measured. This is true especially in the case of U Cephei, SX Cassiopeiae, RZ Scuti and RW Tauri. In all cases a mass of gas is seen projected in front of the receding A-type star, beginning a short time after maximum radial velocity of the latter and preceding the beginning of the partial eclipse. This trend seems to be always in the same direction but its amount differs greatly from system to system. It is very pronounced in U Cephei, but is very little indicated in U Sagittae.

10. In U Sagittae and U Cephei double lines are observed during the partial eclipse, both components of which originate in the ring. By means of a combination of data from several stars we can draw an approximate picture of the flow of gas within the envelope. It seems to follow roughly the outline of a Jacobian equipotential curve in the XY-plane.

11. The amount of matter in a typical ring of this sort can be estimated only very roughly. A value of 10^{-8} times the mass of the brighter component is consistent with the intensities of the emission lines and also with those of the

absorption lines. An average tendency toward expansion amounting to a velocity of less than 1 km./sec. would suffice to exhaust most of the mass of the binary in 10^8 years. If, as seems probable, angular momentum is lost with this mass, it is entirely possible that a system of the type of U Sagittae or even of the type of Plaskett's star could gradually evolve until it would become a typical W Ursae Majoris binary.

If we accept this evolutionary trend it is necessary to explain at one end of the sequence how giant binaries of the kind of Plaskett's star originate and at the other end what ultimately happens to a dwarf binary of the kind of W Ursae Majoris. There is as yet no reliable theory to account for the origin of the massive binaries. But at least we can see that in our galaxy there is no lack of single stars with masses equal to these binaries and with angular momenta sufficient for their creation. The classical fission theory of Darwin and Jeans has now been largely abandoned. But it is not unreasonable to think of a catastrophic process involving the release of nuclear energy which could give rise to the break-up of a rapidly rotating O-type star and the creation of a massive binary.

At the other end of our sequence of close binaries we are confronted by the unmistakable evidence of our observations that there are no single stars of the solar type which have anything like the angular momentum of a W Ursae Majoris system. If the evolutionary trend continues in the direction of bringing the two components together, there must be some mechanism to take care of the excess angular momentum. One way would be a catastrophic explosion with a removal of the star from the H–R diagram and the conversion of its mass into the substratum of the galaxy. We have never observed such a catastrophe, despite the fact that there are many W Ursae Majoris binaries with periods of the order of $\frac{1}{3}$ day and surfaces which are so nearly in contact that it is even now difficult to decide whether they are single stars or pairs of objects.

Another possibility might be the transformation of a W Ursae Majoris system into a slowly rotating red dwarf and a group of attendant planets, the latter carrying with them in their orbits the major portion of the angular momentum of the binary.

The evolutionary ideas presented in this paper are to be regarded as a working hypothesis. I have elsewhere discussed it in greater detail and have presented arguments in its favour and against it. The crucial point of the discussion rests with the recognition that while there is no trouble in regard to the rotational momentum among single and double stars at the upper end of the H–R diagram, there is an unbridged gap at the lower end. The slow rotations of the cool dwarfs would cause us much less trouble if we could believe that these stars are usually accompanied by distant planets so that the total angular momenta of the red dwarfs would not differ drastically from the angular momenta of the early-type stars.

References

(1) O. Struve, *Ap. J.*, **99**, 222, 1944.
(2) A. H. Joy, *Ap. J.*, **71**, 336, 1940.
(3) M. Fowler, *Publ. Allegheny Obs.*, **3**, 11, 1912.
(4) L. Jacchia, *Harvard Obs. Bull.*, No. 915, 33, 1941.
(5) N. Tchudovitchev, *Engelhardt Obs. Bull.*, No. 23, 1946.
(6) R. O. Redman, *M.N.*, **96**, 488, 1936.
(7) F. R. Moulton, *Introduction to Celestial Mechanics*, p. 283, 2nd Ed., New York, 1935.
(8) I am indebted to Dr Lohmann for advance information concerning his work on this subject.

THE SUN'S GENERAL MAGNETIC FIELD

P. A. Sweet

(Communicated by the Director of the Glasgow University Observatory)

(Received 1949 June 7)

Summary

The effects of large-scale convection currents and differential angular velocity on the Sun's general magnetic field are studied in detail, with the following results :—

(i) Cowling's theorem (1) on the dynamo action of convection currents is extended to the case where a small external magnetic field is produced in the convection currents themselves. It is shown that a general magnification of the external field would only be possible if the Hall current were sufficiently large. In the Sun the field so produced would be negligible.

(ii) Ferraro's work (2) on the effect of differential angular velocity on the field is extended to allow for the Hall effect. It is confirmed that the observed differential angular velocity is incompatible with the existence of a steady general magnetic field in the solar atmosphere.

(iii) The concept of transverse conductivity must be used with caution owing to the action of the mechanical force of the resulting magnetic field in setting the gas in motion.

---

1. *Introduction.*—Several attempts have been made in the past to account for the Sun's field by assuming that convection currents act like a dynamo in magnifying some small magnetic field set up initially by a small E.M.F. within the convection currents themselves. Theories on these lines have been put forward by Gurevitch and Lebedinsky (3) in connection with the growth of sunspots, and by Frenkel (4) in connection with the Earth's magnetic field. Neither of these theories proves that the process is electromagnetically possible. As Cowling observes (5), it is difficult to see how new lines of force could be created by dynamo action on an external field, even if the field were produced within the convection currents themselves. Cowling has also shown (5) that, in the absence of Hall current, it would in any case take far too long to build up a general field in this way.

The physical picture which most accurately describes the effect of meridian circulation on the field is not that of a dynamo action at all, but rather that of a fluid in motion between a set of elastic magnetic shells of rotation about the solar axis. If these shells are weak then the fluid distorts them to follow its own lines of flow. On the other hand, if they are sufficiently strong the fluid is constrained to flow between them, so following the lines of magnetic force. In neither case are new lines of force produced by the circulation.

It is more difficult to construct a physical picture to describe the effect of a shearing motion such as arises in differential angular velocity. The distortion of the field increases indefinitely with the magnitude of the shear, whereas in meridian circulation it tends to a finite limit as the velocity of flow increases. A detailed study of this action was given by Ferraro (2) in 1937, neglecting the possible effect of Hall current in the atmosphere.

It is equally difficult to give a physical picture when the Hall effect is not negligible, for it will be shown that meridian circulation does in fact have a dynamo action in this case.

2. *The effect of meridian circulation on the field.*—The most general steady state electromagnetic equations for an axially symmetric magnetic field in the presence of convection currents, symmetrical about the same axis but not necessarily in meridian planes, will first be developed, including the terms due to Hall current. This latter effect was first discussed by Chapman (**6**) in 1928, and developed in detail in the case of the Sun by Cowling (**7**) in 1945.

Instead of the classical equation of conductivity $\boldsymbol{j}=\sigma c^2\boldsymbol{E}$ for a medium of conductivity σ e.m.u., we have

$$\boldsymbol{j}=\sigma^{\mathrm{I}}c^2\boldsymbol{E}_{\perp}+\sigma^{\mathrm{II}}c^2\frac{\boldsymbol{E}_{\perp}\wedge\boldsymbol{H}}{H}+\sigma c^2\boldsymbol{E}_{\parallel}, \tag{2.1}$$

where $\boldsymbol{E}_{\parallel}$ and $\boldsymbol{E}_{\perp}$ are the components of the total electric intensity $\boldsymbol{E}$, parallel and perpendicular, respectively, to the magnetic field $\boldsymbol{H}$, and σ^{I}, σ^{II} and σ are the direct, transverse and normal conductivities respectively. In this work, however, it is more convenient to use the equation of conductivity in the form

$$\boldsymbol{j}=\sigma c^2\left[\frac{\boldsymbol{j}\wedge\boldsymbol{H}}{enc}+\boldsymbol{E}\right], \tag{2.2}$$

where n is the number of free electrons/cm.^{3}, from which the previous equation follows on solving for $\boldsymbol{j}$. If the medium is in motion with velocity $\boldsymbol{v}$ we have

$$\boldsymbol{E}=\frac{\boldsymbol{v}\wedge\boldsymbol{H}}{c}-\operatorname{grad}V+\boldsymbol{E}^{\mathrm{ext}}, \tag{2.3}$$

where $-\operatorname{grad}V$ is the space charge field induced by the motion, and $\boldsymbol{E}^{\mathrm{ext}}$ is the external E.M.F. which produces the external field under distortion. If we put $\boldsymbol{E}^{\mathrm{ext}}=\operatorname{curl}\boldsymbol{H}^{\mathrm{ext}}/4\pi\sigma c$, then $\boldsymbol{H}^{\mathrm{ext}}$ is the field which would normally arise in the absence of convection currents, and may be regarded generally as the field arising from any source, whether it be an inherent magnetic field such as in Blackett's theory (**8**), or a slowly decaying field produced at some previous time. The electromagnetic equation is, as usual,

$$\operatorname{curl}\boldsymbol{H}=4\pi\boldsymbol{j}/c. \tag{2.4}$$

On substituting this expression for $\boldsymbol{j}$ into the equation of conductivity (2.2) we have

$$\operatorname{curl}\boldsymbol{H}=4\pi\sigma\boldsymbol{v}\wedge\boldsymbol{H}-4\pi\sigma c\operatorname{grad}V+4\pi\sigma c\boldsymbol{E}^{\mathrm{ext}}+\frac{4\pi\sigma c}{en}[\operatorname{curl}\boldsymbol{H}\wedge\boldsymbol{H}/4\pi]. \tag{2.5}$$

The last term of this equation is the contribution of the Hall current, and the expression within the bracket is identified as the mechanical force $\boldsymbol{F}^m$ exerted by the magnetic field on the gas. The velocity $\boldsymbol{v}$ is itself dependent to a certain extent on this force, a fact often overlooked in problems where Hall current is important and the gas is presumed to be permanently at rest. When an external E.M.F. is applied to a medium, normally at rest, the magnetic field produced by the transverse current produces a mechanical force of electromagnetic origin which may set up a drift velocity $\boldsymbol{v}$. In certain cases the E.M.F. $\boldsymbol{v}\wedge\boldsymbol{H}/c$ so induced tends to cancel out the Hall effect; the effective conductivity is then the normal conductivity σ, and not that defined by the reduced coefficients σ^{I} and σ^{II}. In the following work it is found that the dependence of $\boldsymbol{v}$ on $\boldsymbol{F}^m$ considerably simplifies the considerations.

In cylindrical polar coordinates (ϖ, ϕ, z) equation (2.5) becomes

$$-\frac{\partial H_\varphi}{\partial z} = 4\pi\sigma(\Omega\varpi H_z - v_z H_\varphi) - 4\pi\sigma c\frac{\partial V}{\partial \varpi} + 4\pi\sigma c E_\omega^{\text{ext}} + \kappa F_\omega^m, \tag{2.6}$$

$$\frac{\partial H_\varpi}{\partial z} - \frac{\partial H_z}{\partial \varpi} = 4\pi\sigma(v_z H_\varpi - v_\varpi H_z) + 4\pi\sigma c E_\varphi^{\text{ext}} + \kappa F_\varphi^m, \tag{2.7}$$

$$\frac{1}{\varpi}\frac{\partial}{\partial \varpi}(\varpi H_\varphi) = 4\pi\sigma(v_\varpi H_\varphi - \Omega\varpi H_\varpi) - 4\pi\sigma c\frac{\partial V}{\partial z} + 4\pi\sigma c E_z^{\text{ext}} + \kappa F_z^m, \tag{2.8}$$

where $\kappa = 4\pi\sigma c/en$, v_ϖ and v_z are the meridian components of the circulation velocity $\boldsymbol{v}$, and Ω is the angular velocity of the gas about the solar axis. Since the magnetic field is solenoidal and symmetrical about the axis we can write

$$H_\varpi = -\frac{1}{\varpi}\frac{\partial B}{\partial z} \quad \text{and} \quad H_z = +\frac{1}{\varpi}\frac{\partial B}{\partial \varpi},$$

where $B = B(\varpi, z)$ is a stream function, constant on the lines of magnetic force. Since the field has no singularities, the lines $B = \text{const.}$ in a given meridian plane form a nest, or series of nests, of closed curves C in that plane. These nests fall into two categories, those containing a single limiting point, or pole, P, which will be called polar nests, and those enveloping two or more polar nests as shown in Fig. 1. The bounding curve separating the polar nests N_1 and N_2 from the enveloping nest N_{12} has a singularity at the point P_{12}, which is a neutral point of the meridian component $\boldsymbol{H}^m$ of the field.

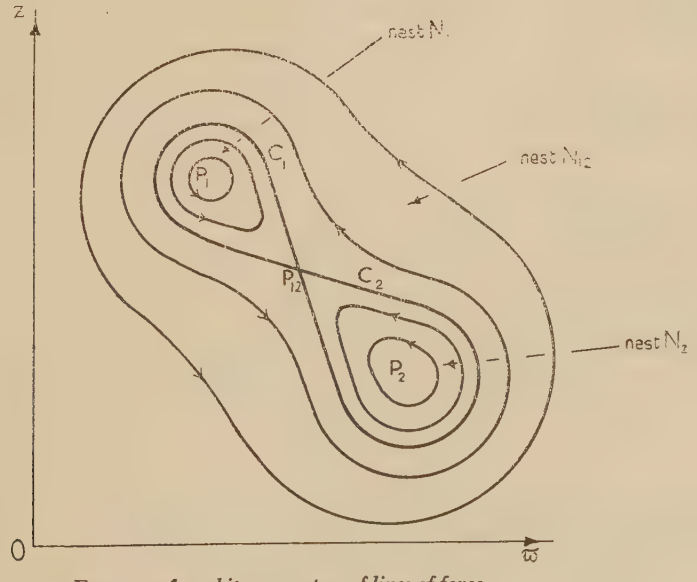

FIG. 1.—*An arbitrary system of lines of force.*

On substituting for H_ϖ and H_z in terms of B in (2.7) we derive the following equation for B:—

$$\frac{\partial}{\partial \varpi}\left(\frac{1}{\varpi}\frac{\partial B}{\partial \varpi}\right) + \frac{1}{\varpi}\frac{\partial^2 B}{\partial z^2} = \frac{4\pi\sigma}{\varpi}\left(v_\varpi\frac{\partial B}{\partial \varpi} + v_z\frac{\partial B}{\partial z}\right) - 4\pi\sigma c E_\varphi^{\text{ext}} - \kappa F_\varphi^m. \tag{2.9}$$

If the space charge function V is eliminated between equations (2.6) and (2.8) a similar equation for H_φ is obtained, thus:—

$$\frac{\partial}{\partial \varpi}\left[\frac{1}{\sigma\varpi}\frac{\partial}{\partial \varpi}(\varpi H_\varphi)\right]+\frac{\partial}{\partial z}\left[\frac{1}{\sigma}\frac{\partial H_\varphi}{\partial z}\right]$$
$$=4\pi\left[\frac{\partial}{\partial \varpi}(v_\varpi H_\varphi)+\frac{\partial}{\partial z}(v_z H_\varphi)\right]-4\pi\left[\frac{\partial}{\partial \varpi}(\Omega\varpi H_\varpi)+\frac{\partial}{\partial z}(\Omega\varpi H_z)\right]$$
$$+4\pi c\left[\frac{\partial E_z^{\text{ext}}}{\partial \varpi}-\frac{\partial E_\varpi^{\text{ext}}}{\partial z}\right]+\left[\frac{\partial}{\partial \varpi}\left(\frac{\kappa}{\sigma}F_z^m\right)-\frac{\partial}{\partial z}\left(\frac{\kappa}{\sigma}F_\varpi^m\right)\right]. \qquad (2.10)$$

Now the equation of continuity of the gas motion is

$$\frac{1}{\varpi}\frac{\partial}{\partial \varpi}(\rho\varpi v_\varpi)+\frac{\partial}{\partial z}(\rho v_z)=0, \qquad (2.11)$$

hence, on putting $H_\varphi=\rho\varpi\bar{B}$, equation (2.10) can be written, after a little reduction,

$$\frac{\sigma}{\rho\varpi^2}\left\{\frac{\partial}{\partial \varpi}\left[\frac{1}{\sigma\varpi}\frac{\partial}{\partial \varpi}(\rho\varpi^2\bar{B})\right]+\frac{\partial}{\partial z}\left[\frac{\varpi}{\sigma}\frac{\partial}{\partial z}(\rho\bar{B})\right]\right\}$$
$$=\frac{4\pi\sigma}{\varpi}\left[v_\varpi\frac{\partial\bar{B}}{\partial \varpi}+v_z\frac{\partial\bar{B}}{\partial z}\right]-\frac{4\pi\sigma}{\rho\varpi}\left[H_\varpi\frac{\partial\Omega}{\partial \varpi}+H_z\frac{\partial\Omega}{\partial z}\right]$$
$$+\frac{4\pi\sigma c}{\rho\varpi^2}\left[\frac{\partial E_z^{\text{ext}}}{\partial \varpi}-\frac{\partial E_\varpi^{\text{ext}}}{\partial z}\right]+\frac{\sigma}{\rho\varpi^2}\left[\frac{\partial}{\partial \varpi}\left(\frac{\kappa}{\sigma}F_z^m\right)-\frac{\partial}{\partial z}\left(\frac{\kappa}{\sigma}F_\varpi^m\right)\right]. \qquad (2.12)$$

The first term on the right-hand side of this equation is similar to the corresponding term of equation (2.9). As these are the only terms containing the meridian component $\boldsymbol{v}^m$ of $\boldsymbol{v}$, $\boldsymbol{v}^m$ will be expected therefore to have similar effects on both functions B and $\bar{B}$. The equations (2.9) and (2.12) will first be considered, neglecting the Hall current. The last term on the right-hand side of both equations vanishes in this case, and we are left to consider the equations

$$\frac{\partial}{\partial \varpi}\left(\frac{1}{\varpi}\frac{\partial B}{\partial \varpi}\right)+\frac{1}{\varpi}\frac{\partial^2 B}{\partial z^2}=\frac{4\pi\sigma}{\varpi}\left(v_\varpi\frac{\partial B}{\partial \varpi}+v_z\frac{\partial B}{\partial z}\right)-4\pi\sigma c E_\varphi^{\text{ext}}, \qquad (2.13)$$

$$\frac{\sigma}{\rho\varpi^2}\left\{\frac{\partial}{\partial \varpi}\left[\frac{1}{\sigma\varpi}\frac{\partial}{\partial \varpi}(\rho\varpi^2\bar{B})\right]+\frac{\partial}{\partial z}\left[\frac{\varpi}{\sigma}\frac{\partial}{\partial z}(\rho\bar{B})\right]\right\}$$
$$=\frac{4\pi\sigma}{\varpi}\left[v_\varpi\frac{\partial\bar{B}}{\partial \varpi}+v_z\frac{\partial\bar{B}}{\partial z}\right]-\frac{4\pi\sigma}{\rho\varpi}\left[H_\varpi\frac{\partial\Omega}{\partial \varpi}+H_z\frac{\partial\Omega}{\partial z}\right]$$
$$+\frac{4\pi\sigma c}{\rho\varpi^2}\left[\frac{\partial E_z^{\text{ext}}}{\partial \varpi}-\frac{\partial E_\varpi^{\text{ext}}}{\partial z}\right]. \qquad (2.14)$$

We shall first consider the equation for B which determines the meridian field as distorted by the meridian circulation. Equation (2.13) is independent of the transverse component H_φ of the field, and also of the angular velocity Ω. Before proceeding further, there is one interesting point to be noticed. If there is no external field, then $E_\varphi^{\text{ext}}=0$; now B is constant on any line of force of the meridian component $\boldsymbol{H}^m$ of the field, hence in this case $B=$ const. everywhere is a solution in the region enclosed by any given line of force. But the regular solution of such an elliptic linear partial differential equation of the second order in two independent variables is determined uniquely within any closed curve by its values on that

curve (9). The above solution is therefore unique, hence $\boldsymbol{H}^m$ vanishes in the region enclosed by any given line of force, and therefore vanishes everywhere. This is an alternative proof of Cowling's theorem (1) that dynamo currents are unable to maintain a magnetic field in the absence of an external field. So far no restrictions have been imposed on the meridian flow $\boldsymbol{v}^m$ in this equation; the theorem therefore holds even if the flow does not satisfy the equation (2.11). Such a case may arise in a quasi-steady state where material flows steadily into a given region from all directions. To continue, it will now be shown that the solution of (2.13) tends to a finite limit as $|\boldsymbol{v}^m| \to \infty$, subject to the equation of continuity (2.11), and with given lines of flow. This limiting solution will be found explicitly.

On multiplying (2.13) by $\rho\varpi^2/\sigma$ and integrating over an area S_{12} between two closed curves C_1 and C_2 of magnetic force, we obtain the equation

$$\iint_{S_{12}} \frac{\rho\varpi^2}{\sigma}\left[\frac{\partial}{\partial\varpi}\left(\frac{1}{\varpi}\frac{\partial B}{\partial\varpi}\right)+\frac{1}{\varpi}\frac{\partial^2 B}{\partial z^2}\right] dS$$
$$= 4\pi \iint_{S_{12}} \rho\varpi\left[v_\varpi \frac{\partial B}{\partial\varpi}+v_z\frac{\partial B}{\partial z}\right] dS - 4\pi c \iint_{S_{12}} \rho\varpi^2 E_\varphi^{\text{ext}}\, dS. \qquad (2.15)$$

The integral containing the velocity components vanishes, for we have

$$\iint_{S_{12}} \rho\varpi\left[v_\varpi\frac{\partial B}{\partial\varpi}+v_z\frac{\partial B}{\partial z}\right] dS = \iint_{S_{12}}\left[\frac{\partial}{\partial\varpi}(\rho\varpi v_\varpi B)+\frac{\partial}{\partial z}(\rho\varpi v_z B)\right] dS$$
$$- \iint_{S_{12}} B\left[\frac{\partial}{\partial\varpi}(\rho\varpi v_\varpi)+\frac{\partial}{\partial z}(\rho\varpi v_z)\right] dS. \qquad (2.16)$$

The first integral on the right-hand side of this equation may be transformed by Stokes' theorem into the line integrals

$$\int_{C_2} B\rho\varpi(v_\varpi dz - v_z d\varpi) - \int_{C_1} B\rho\varpi(v_\varpi dz - v_z d\varpi).$$

On noting that B is constant on lines of force, and transforming back into surface integrals, this expression may be written

$$B_2\iint_{S_2}\left[\frac{\partial}{\partial\varpi}(\rho\varpi v_\varpi)+\frac{\partial}{\partial z}(\rho\varpi v_z)\right] dS - B_1\iint_{S_1}\left[\frac{\partial}{\partial\varpi}(\rho\varpi v_\varpi)+\frac{\partial}{\partial z}(\rho\varpi v_z)\right] dS,$$

where S_1 and S_2 are the whole areas enclosed by the curves C_1 and C_2 respectively, and B_1 and B_2 are the values of B on these curves. The integrands of these expressions vanish everywhere on account of the equation of continuity. The second integral on the right-hand side of (2.16) vanishes similarly, hence equation (2.15) reduces to

$$\iint_{S_{12}} \frac{\rho\varpi^2}{\sigma}\left[\frac{\partial}{\partial\varpi}\left(\frac{1}{\varpi}\frac{\partial B}{\partial\varpi}\right)+\frac{1}{\varpi}\frac{\partial^2 B}{\partial z^2}\right] dS = -4\pi c\iint_{S_{12}} \rho\varpi^2 E_\varphi^{\text{ext}}\, dS. \qquad (2.17)$$

Except for the shape of the region S_{12} this equation is independent of the meridian circulation $\boldsymbol{v}^m$. We shall now derive an expression for B in terms of the shape of the lines of force.

The whole domain $\varpi \geqslant 0$ of the equation can be covered by one or more nests of closed lines of force. Within each nest an orthogonal coordinate system (u, α) can be constructed, $u(\varpi, z) = \text{const.}$ and $\alpha(\varpi, z) = \text{const.}$ defining the lines of force and their orthogonal trajectories respectively. u may be defined as the distance along a particular orthogonal trajectory Γ, measured from the inner boundary of the nest. For a polar nest this would be the pole of the nest. For an enveloping nest, α may be defined as the distance along the inner boundary of the nest measured

in an anticlockwise direction from its point of intersection with Γ. For a polar nest, α is taken as the angle between the relevant orthogonal trajectory and the trajectory Γ, at the pole. In the case of a nest consisting of a set of concentric circles, (u, α) are simply the polar coordinates (r, θ). Hence for each nest u ranges from zero to some finite value u', finite even for the outermost nest, since in this case all the orthogonal trajectories, except one limiting one, terminate on the z-axis, cutting it at right angles. For polar nests, α ranges from zero to 2π, and for enveloping nests, from zero to some finite positive value α_1.

In this problem each set (u, α) are functions of ϖ and z determined by the meridian circulation $\boldsymbol{v}^m$, and the E.M.F. E_φ^{ext}, but for the moment they will be regarded as known functions. An explicit solution for $\boldsymbol{H}^m$ is then obtainable in terms of (u, α) as follows: consider a narrow strip S_{12} within a particular nest N_1, bounded by two adjacent curves $u = u_0$, $u = u_1$ of N_1. On transforming the integrals in equation (2.17) into integrals with respect to the variables (u, α), and noting that $B = B(u)$, it follows that

$$\int_{u_0}^{u_1}\int_0^{\alpha_1}\frac{\rho\varpi}{\sigma}\left[\frac{d^2B}{du^2}\left\{\left(\frac{\partial u}{\partial \varpi}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2\right\}+\frac{dB}{du}\left\{\frac{\partial^2 u}{\partial \varpi^2}+\frac{\partial^2 u}{\partial z^2}-\frac{1}{\varpi}\frac{\partial u}{\partial \varpi}\right\}\right]J\,d\alpha\,du$$
$$=-4\pi c\int_{u_0}^{u_1}\int_0^{\alpha_1}\rho\varpi^2 E_\varphi^{\text{ext}} J\,d\alpha\,du, \tag{2.18}$$

where $J = \partial(\varpi, z)/\partial(u, \alpha)$. On proceeding to the limit $u_1 \to u_0$, and dropping the suffix from u_0, we have

$$\frac{d^2B}{du^2}\int_0^{\alpha_1}\frac{\rho\varpi}{\sigma}\left[\left(\frac{\partial u}{\partial \varpi}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2\right]J\,d\alpha+\frac{dB}{du}\int_0^{\alpha_1}\frac{\rho\varpi}{\sigma}\left[\frac{\partial^2 u}{\partial \varpi^2}+\frac{\partial^2 u}{\partial z^2}-\frac{1}{\varpi}\frac{\partial u}{\partial \varpi}\right]J\,d\alpha$$
$$=-4\pi c\int_0^{\alpha_1}\rho\varpi^2 E_\varphi^{\text{ext}} J\,d\alpha. \tag{2.19}$$

The three integrals in this equation are known functions of u. On putting $dB/du = G$ we have, finally,

$$H_\varpi = -\frac{1}{\varpi}\frac{\partial u}{\partial z}G \qquad \text{and} \qquad H_z = \frac{1}{\varpi}\frac{\partial u}{\partial \varpi}G,$$

where G is given by the first order ordinary differential equation

$$U_1(u)\frac{dG}{du}+U_2(u)G+U_3(u)=0. \tag{2.20}$$

Here

$$U_1=\int_0^{\alpha_1}\frac{\rho\varpi}{\sigma}\left[\left(\frac{\partial u}{\partial \varpi}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2\right]J\,d\alpha, \qquad U_2=\int_0^{\alpha_1}\frac{\rho\varpi}{\sigma}\left[\frac{\partial^2 u}{\partial \varpi^2}+\frac{\partial^2 u}{\partial z^2}-\frac{1}{\varpi}\frac{\partial u}{\partial \varpi}\right]J\,d\alpha$$

and

$$U_3=\int_0^{\alpha_1}4\pi c\rho\varpi^2 E_\varphi^{\text{ext}} J\,d\alpha.$$

On solving (2.20) we obtain

$$G=-\exp\left[-\int^u (U_2/U_1)\,du\right]\int_{\bar{u}}^{u}\left\{\exp\left[\int^u (U_2/U_1)\,du\right]U_3/U_1\right\}du, \tag{2.21}$$

where $\bar{u}$ is some value of u determined by the initial condition for G. This shows that, provided no singularities other than a finite number of double points develop in the lines of force, the field remains bounded as $|\boldsymbol{v}^m| \to \infty$ for given lines of flow. Moreover the field is of the same order of magnitude as the field which would be produced by the E.M.F. E_φ^{ext} in the absence of meridian circulation.

In a rigorous treatment the assumption about the singularities in the shape of the lines of force would have to be proved independently. This will not be attempted in this paper. However, if the field were in fact bounded as $|\boldsymbol{v}^m| \to \infty$, then the term $\boldsymbol{v}^m \,.\, \mathrm{grad}\, B$ on the right-hand side of equation (2.13) would also be bounded. Hence $\boldsymbol{H}^m$ would tend to become parallel to $\boldsymbol{v}^m$; the lines of force would thus tend into coincidence with the lines of flow, which partly justifies the assumption.

It remains to determine the initial condition for G. For a polar nest we have $G(0) = 0$ as the initial condition, since otherwise $\boldsymbol{H}^m$ would have a singularity at the pole. In this case $\bar{u} = 0$, and the resulting solution for a polar nest is therefore

$$G = -\exp\left[-\int^u (U_2/U_1)\,du\right]\int_0^u \left\{\exp\left[\int^u (U_2/U_1)\,du\right] U_3/U_1\right\} du. \quad (2.25)$$

If this is the only nest of the system, we note that, since $G(u')$ is finite and $\partial u/\partial \varpi$ and $\partial u/\partial z$ vanish at infinity, $\boldsymbol{H}^m$ also vanishes at infinity as required.

The source of the field is the applied E.M.F. E_φ^{ext}, which appears in the function U_3. It follows from (2.25) that, if there is no E.M.F. within a line of flow $C(u_0)$, say, then $U_3 = 0$ for $u \leqslant u_0$, hence $G = 0$ for $u \leqslant u_0$, and there is therefore no meridian magnetic field within $C(u_0)$. This establishes the important result that a meridian circulation of sufficient velocity round a given region screens the region from an external magnetic field.

For an enveloping nest, such as N_{12} in Fig. 1, the initial condition is in general no longer $G(0) = 0$, but is a condition which fits the nest on to the interior nests N_1 and N_2. For finite values of $\boldsymbol{v}^m$ the initial value of G is equal to the known value on the outer boundaries C_1 and C_2 of N_1 and N_2. At first sight it may seem that the values of G on C_1 and C_2, as found from the two solutions of the type (2.25) for N_1 and N_2, would differ in general. It must be remembered, however, that the configurations of the lines of force and the magnitudes of the field in each nest are not independent, hence the two solutions are not independent. In the limit as $|\boldsymbol{v}^m| \to \infty$, the field $\boldsymbol{H}^m$ in general develops a discontinuity on C_1 and C_2. The initial condition for the field in N_{12} in this case is given by

$$\mathrm{Lt}_{C \to C_1 + C_2}\left[\int_C \frac{\rho\varpi^2}{\sigma}\boldsymbol{H}_0^m \,.\, d\boldsymbol{s}\right] = \int_{C_1} \frac{\rho\varpi^2}{\sigma}\boldsymbol{H}_1^m \,.\, d\boldsymbol{s} + \int_{C_2} \frac{\rho\varpi^2}{\sigma}\boldsymbol{H}_2^m \,.\, d\boldsymbol{s}, \quad (2.26)$$

where C is a curve of N_{12}, and $\boldsymbol{H}_0^m$, $\boldsymbol{H}_1^m$ and $\boldsymbol{H}_2^m$ are the limiting fields $\boldsymbol{H}^m$, as $|\boldsymbol{v}^m| \to \infty$ in the nests N_{12}, N_1 and N_2 respectively. To prove this condition consider equation (2.17) and take the region S_{12} as the area between the curves C, C_1 and C_2. On rearranging the integrand on the left-hand side of (2.17) and transforming part of it into line integrals by Stokes' theorem, we have

$$\begin{aligned}\int_C \frac{\rho\varpi^2}{\sigma}\boldsymbol{H}_0^m \,.\, d\boldsymbol{s} - \int_{C_1} \frac{\rho\varpi^2}{\sigma}\boldsymbol{H}_1^m \,.\, d\boldsymbol{s} - \int_{C_2} \frac{\rho\varpi^2}{\sigma}\boldsymbol{H}_2^m \,.\, d\boldsymbol{s} \\ + \iint_{S_{12}} (B - B_0) K\,dS + (B_1 - B_0)\iint_{S_1} K\,dS + (B_2 - B_0)\iint_{S_2} K\,dS \\ = -4\pi c \iint_{S_{12}} \rho\varpi^2 E_\varphi^{\mathrm{ext}}\,dS, \quad (2.27)\end{aligned}$$

where B_0, B_1 and B_2 are the values of B on C, C_1 and C_2 respectively, S_1 and S_2 are the areas enclosed by the curves C_1 and C_2 respectively, and

$$K = \varpi\left[\frac{1}{\varpi}\frac{\partial}{\partial\varpi}\left\{\frac{1}{\varpi}\frac{\partial}{\partial\varpi}\left(\frac{\rho\varpi^2}{\sigma}\right)\right\} + \frac{\partial^2}{\partial z^2}\left(\frac{\rho}{\sigma}\right)\right].$$

Now $\boldsymbol{H}^m$ is bounded as $|\boldsymbol{v}^m| \to \infty$, hence the limiting function B is continuous. Thus $|B_1 - B_0|$ and $|B_2 - B_0|$ can be made as small as desired by taking C sufficiently close to the limiting curve $C' = C_1 + C_2$. ($B_1 = B_2$ since the nests N_1 and N_2 have a common point P_{12}.) The last three terms on the left-hand side of (2.27) can therefore be made arbitrarily small. The integral on the right-hand side of (2.27) can also be made arbitrarily small, since it does not contain B in any case. The boundary condition (2.26) follows on taking equation (2.27) to the limit $C \to C'$.

In practical cases, of course, $|\boldsymbol{v}^m|$ is limited, but it can be seen by inspection of the electromagnetic equation that the limiting case $|\boldsymbol{v}^m| \to \infty$ is a good approximation when $|\boldsymbol{v}^m| \gg 1/4\pi\sigma d$, where d is a typical distance of variation of the direction of flow in the neighbourhood. In the solar interior, taking typical conditions $\sigma \sim 10^{-5}$ e.m.u. and $d \sim 10^{10}$ cm., the above value is no greater than 10^{-6} cm./sec. If $|\boldsymbol{v}^m|$ is anywhere observed to rise appreciably above this value, then the lines of force there follow the lines of flow. In these circumstances, either the forces producing the flow have been strong enough to distort the lines of force to follow the lines of flow, or else the magnetic field itself has been strong enough to direct the flow along the lines of force. In either case the actual drift of the material across the lines of force is limited to a value appreciably below this critical value. In the reversing layer, where $\sigma \sim 10^{-9}$ e.m.u., for a drift velocity in the surface maintained over an appreciable range of latitude, d can again be taken as $O(10^{10})$ cm., giving a critical value of the order of 10^{-2} cm./sec., a negligible velocity.

3. *The effect of differential angular velocity on the field.*—Consider the transverse component H_φ of the field and its associated function $\bar{B}$ as determined by equation (2.14), again neglecting the Hall current. The second and third terms on the right-hand side of this equation are independent of $\boldsymbol{v}^m$ for a given meridian component $\boldsymbol{H}^m$, and may be regarded as finite known functions. The analysis of the previous section may now be applied to the equation, with the result that, as $|\boldsymbol{v}^m|$ increases, $\bar{B}$ remains finite and of the same order of magnitude as in the absence of meridian circulation, and the lines $\bar{B} = \text{const.}$ tend to become distorted by the meridian circulation to follow the lines of flow. The second term on the right-hand side of the equation is due to differential angular velocity, the effect of which on the magnetic field will now be examined as in Ferraro's work (**2**). The order of magnitude of the effect of an arbitrary variation of angular velocity is easily found by comparing the left-hand side of equation (2.14) with the second term on its right-hand side. This shows that in general $H_\varphi = O(4\pi\sigma d^2 H \Delta\Omega)$, where d is a typical distance of variation of Ω, $H = |\boldsymbol{H}^m|$, and $\Delta\Omega$ is a typical variation in angular velocity Ω.

In the Sun Ω is observed to vary, at least on the surface, by about 20 per cent, hence taking $\sigma \sim 10^{-9}$ e.m.u., $H \sim 50$ gauss and $\Delta\Omega \sim 6 \times 10^{-7}$ radians/sec. as conditions typical of the reversing layer, together with $d \sim 10^{10}$ cm., we would have $H_\varphi \sim 4 \times 10^7$ gauss, an impossibly high field. Such an observed differential angular velocity is therefore only possible if it occurs in such a way that the term

$$H_\varpi \frac{\partial \Omega}{\partial \varpi} + H_z \frac{\partial \Omega}{\partial z} = \boldsymbol{H}^m . \operatorname{grad} \Omega$$

is small, i. e. if the direction of variation of Ω is perpendicular to the meridian magnetic field. Hence in the presence of a magnetic field strong enough to be observable, the Sun could only rotate independently in shells along the lines of force of the field, the angular velocity being constant along any given line of force.

The shells of magnetic force can be regarded as being extremely resistant to distortion by transverse forces, for any force tending to produce such distortion very quickly produces a large transverse component H_φ with a resultant large restoring component F_φ^m of the electromagnetic force. In steady conditions, therefore, H_φ is determined not by equation (2.10) but by the solar mechanical forces available for maintaining the resulting component F_φ^m of the electromagnetic force. If the effect of viscosity can be neglected, the only important possible force available is the Coriolis force on convection currents flowing in meridian planes. Apart from this force, we would have $H_\varphi = 0$ in equation (2.14). In this case (2.14) reduces to

$$\boldsymbol{H}^m \,.\, \text{grad}\,\Omega = \frac{c}{\varpi}\left[\frac{\partial E_z^{\text{ext}}}{\partial \varpi} - \frac{\partial E_\omega^{\text{ext}}}{\partial z}\right]. \tag{3.1}$$

Hence an E.M.F. in meridian planes is capable of producing a small differential angular velocity. However, the E.M.F.s likely to arise under steady conditions in the Sun are far too small to produce a noticeable differential angular velocity if the meridian component of the field is large enough to be observed.

4. *The effect of Hall current on the meridian component of the field.*—The term κF_φ^m in equation (2.7) contains H_φ, and so couples this equation with the equation (2.12) for the transverse component. In steady conditions F_φ^m is in equilibrium with the Coriolis force on any meridian circulation currents present, hence

$$F_\varphi^m = 2\rho\Omega v_\varpi.$$

Thus $\kappa F_\varphi^m = 8\pi\sigma c\rho\Omega v_\varpi/en \sim 8 \times 10^{-9}\sigma v_\varpi$ in the chromosphere, where ρ/n is taken as approximately the mass m_H of a hydrogen atom. In the reversing layer it is of the order of magnitude $8 \times 10^{-6}\sigma v_\varpi$, where ρ/n is taken as approximately $1000 m_H$ owing to the low ionization there. The term is therefore likely to have its greatest effect in the reversing layer.

According to an estimate made by Eddington (**10**), the steady meridian circulation velocity cannot greatly exceed a value of the order of 2×10^{-4} cm./sec., hence in the reversing layer $\kappa F_\varphi^m = O(10^{-18})$. The typical distance d of variation of the field as affected by this term is obtained roughly by comparing it with the resulting order of magnitude H/d of the left-hand side of (2.7). With $H \sim 50$ gauss this would give $d = O(10^{19})$ cm., which shows that the effect of the Hall current is negligible.

5. *The effect of Hall current on the transverse component.*—It has been shown that, in steady conditions, this component is determined not by its electromagnetic equation, but by the external mechanical forces. The Hall current therefore has no effect on it.

However, the relevant electromagnetic equation (2.12) must still be satisfied. The order of magnitude of F_φ^m shows that if the meridian component of the field is large enough to be observed, then H_φ is small compared with it. Hence, if the field extends into the atmosphere, we can neglect those terms in (2.12) which contain $\overline{B}$. The equation then becomes an equation determining, not H_φ, but the differential angular velocity $\partial\Omega/\partial\varpi$ and $\partial\Omega/\partial z$ in terms of the meridian components E_z^{ext}, E_ϖ^{ext} of the E.M.F. $\boldsymbol{E}^{\text{ext}}$ and the meridian field $\boldsymbol{H}^m$. The angular velocity is therefore given by

$$\boldsymbol{H}^m \,.\, \text{grad}\,\Omega = -\frac{c}{\varpi}\{\text{curl}\,\boldsymbol{E}^{\text{ext}}\}_\varphi + \frac{c}{\varpi}\left[\frac{\partial}{\partial\varpi}\left(\frac{F_z'}{en}\right) - \frac{\partial}{\partial z}\left(\frac{F_\varpi'}{en}\right)\right], \tag{5.1}$$

where F_ϖ', F_z' are the meridian components of the electromagnetic mechanical force, neglecting H_φ. It has already been mentioned that the first term on the

right-hand side is not likely to contribute appreciably to the equation. The second term on the right-hand side is in general of the order of magnitude $H^2\kappa/(4\pi)^2\sigma d^3$, where d is a typical length. Hence the order of magnitude of the differential angular velocity produced by Hall current is given by

$$\Delta\Omega \sim H\kappa/(4\pi)^2\sigma d^2 = cH/4\pi end^2 \text{ radians/sec.}$$

If the field extended into the chromosphere where $n \sim 10^{11}$, we would have $\Delta\Omega \sim 10^{-11}$ on taking $H = 50$ gauss, and $d \sim 10^{10}$ cm. The effect on the angular velocity is therefore negligible.

6. *Conclusions.*—(i) The theory disposes of the possibility of a steady general magnetic field, large enough to be observed, arising from dynamo effects. The work of Section 4, however, shows that the Hall effect would make a dynamo action possible under certain conditions not obtaining in the Sun.

(ii) The lines of magnetic force behave like rigid shells even in the chromosphere, where the Hall current might be expected to have some effect on them.

The equatorial acceleration in angular velocity would be effected by the inner shells rotating faster than the outer shells. If the field were limited in the reversing layer, the lines of force would be flattened to run in the surface, and although the equatorial acceleration could still be effected by the inner shells rotating faster than the outer shells, there would also be a sharp decrease in angular velocity with height in the reversing layer. This, however, is not observed. On the other hand, if the field extended into the chromosphere, then this would have almost the same angular velocity as the reversing layer. This again is contrary to observation, for observations of the $H\alpha$ line indicate that the chromosphere rotates faster and more uniformly over the whole disk.

There can therefore be no steady general magnetic field in the solar atmosphere, if the observations of differential angular velocity have been interpreted correctly.

(iii) If there were an appreciable magnetic field in the solar atmosphere, the gas would act, at least in the cases discussed, as if it had the normal conductivity σ, as in the absence of a magnetic field. In any other application of the equation of conductivity in its usual form (2.1), care should be taken that the magnetic field tending to arise from Hall current does not react on the equation by setting the gas in motion, so introducing into it an additional E.M.F. $\boldsymbol{v}\wedge\boldsymbol{H}/c$. A safer approach to electromagnetic problems in the solar atmosphere would be along the lines given in Section 2 in developing the electromagnetic equation (2.5).

University Observatory,
Glasgow:
1949 *June* 6.

References

(1) Cowling, T. G., *M.N.*, **94,** 39, 1933.
(2) Ferraro, V. C. A., *M.N.*, **97,** 458, 1937.
(3) Gurevitch, L. E. and Lebedinsky, A. I., *C.R. Acad. Sci., U.S.S.R.*, **49,** 92, 1945.
(4) Frenkel, J., *C.R. Acad. Sci., U.S.S.R.*, **49,** 98, 1945.
(5) Cowling, T. G., *M.N.*, **105,** 166, 1945.
(6) Chapman, S., *M.N.*, **89,** 54, 1928.
(7) Cowling, T. G., *Proc. Roy. Soc.* A, **183,** 453, 1945.
(8) Blackett, P. M. S., *Nature*, **159,** 658, 1947.
(9) Courant, R. and Hilbert, D., *Die Grundlehren d. math. Wiss.*, **48,** 275, 1937.
(10) Eddington, A. S., *M.N.*, **90,** 54, 1929.

THE LUMINOSITY-VELOCITY RELATION FOR PULSATING STARS

E. A. Milne, F.R.S.

(Received 1949 June 3)

Summary

It is shown by means of an application of the theorem of the virial that for a pulsating star the variation in luminosity should be strictly in phase with the surface-velocity (as observed for Cepheids), provided the star is sufficiently homogeneous; and that the variation in luminosity should be opposite in phase to the surface-velocity (as observed for long-period variables), provided there is a sufficient mass-concentration towards the centre. It is not assumed that the pulsations are adiabatic. It is pointed out that strictly adiabatic pulsations should give zero luminosity-variation, and this is confirmed by substituting the period-formula of Ledoux and Pekeris, who assumed adiabatic pulsations.

---

1. *Introduction.*—A few years before the war I attempted in an unpublished investigation to apply the theorem of the virial to the radial oscillations of a star, but desisted owing to a certain difficulty. Since then, quite independently, the theorem of the virial has been successfully applied by Ledoux (**5**) to obtain a proof of a formula, for the period of oscillation, obtained earlier by another method by himself and Pekeris (**6**). I had overlooked these investigations at the time, but the appearance of Rosseland's new monograph (**11**) has prompted me to reconsider my earlier work. I am also much indebted to discussions with Mr C. Prasad.

In the following brief paper I apply the theorem of the virial, in a way differing from Ledoux's, to obtain a formula for the variation in luminosity of a pulsating star under certain assumptions, but *without* making the assumption that the pulsations are adiabatic. The formula involves the outward velocity of the star's surface, and it gives at once that the increase in luminosity δL at any moment is proportional to the outward velocity of the star's surface, $\dot{r}_0$, to the first order; and that the coefficient of proportionality is positive, provided a certain inequality is satisfied, an inequality depending only on the relative mass-distribution in the interior. This inequality will be likely to be reversed when we pass to long-period variables. Thus the theory gives the observed agreement of maximum velocity of expansion with maximum luminosity, provided the mass-distribution is appropriate, and reverses the sign of the agreement for variables of long period. This removes a fundamental objection to the pulsation theory, as hitherto developed, for previous investigations on the pulsation theory have always given maximum luminosity as occurring at *minimum radius*. Attempts to build up the missing quarter-period of phase from other considerations have been on the whole unsuccessful, and in any case it has always seemed unlikely that so clean-cut and characteristic a property of Cepheid variation could arise from an accumulation of small phase-differences.

But now a most important point arises. When in my formula for δL in terms of $\dot{r}_0$ I substitute the formula of Ledoux and Pekeris for the period, I find that δL vanishes. The reason for this is not far to seek. Both Ledoux and Pekeris, and

Ledoux again in his virial method of deriving the formula for the period, assumed that the oscillations were *adiabatic*, following the lead of Eddington. This is equivalent to assuming that the star's material can be considered as partitioned into adiabatically-walled cells. But in that case there can be no flow of radiant energy during the pulsation. Since the flow of radiation is inhibited by the hypothetical adiabatic cell-walls, there can be no fluctuations of luminosity during the cycle. The impervious, totally reflecting cell-walls would indeed preclude any flow of radiation at all, and so would give a configuration of zero-luminosity! For even the normal interchange of radiation between element and element, which gives rise to the observed luminosity, would be inhibited.

The assumption that in a pulsating star the oscillations are adiabatic was originally made by Eddington (**2**), but I have long considered this assumption as essentially contradicting the other assumptions on which the equilibrium theory of stellar structure is built up. Indeed, I stated this difficulty in my recent review (**7**) of Rosseland's monograph, a review written before the results of the present paper were obtained.

Thus the formula of Ledoux and Pekeris for the period is rigorously true on the adiabatic hypothesis (to the first order of small quantities and on certain minor other assumptions): but this hypothesis automatically makes luminosity fluctuations impossible. Their formula for the period is of course *very nearly* true to nature, and this is borne out by the present analysis, for unless it were very nearly true, the variation in luminosity given by the present theory would be far too large.

The simple considerations of this paper throw a flood of light on the failure of previous investigators to account for the phase-relation in Cepheid variation on the pulsation theory. As remarked above, it has for some time been evident that the exactness of the agreement between the time of maximum luminosity and the time of maximum velocity of approach of the star's surface made it improbable that the necessary phase-lag given on the usual theory could be removed by effects in the outer layers, such as the existence of a convective zone below the surface; the agreement must depend on some fundamental property of the oscillations which is not covered by the usual theory. We now see the Achilles' heel of the usual theory, namely its assumption of adiabatic oscillations, for these imply no change of luminosity at all. Hence such a refinement as the phase-agreement between maximum luminosity and maximum velocity of expansion must necessarily escape the usual theory altogether.

2. *Application of the virial theorem.*—Owing to the fundamental nature of the phase-relation in Cepheids between luminosity and velocity, it is desirable to establish it in the first instance on as simple and broad a basis as possible. With this object I consider a gravitating set of particles arranged with spherical symmetry, and in the first instance I neglect radiation pressure. The equation of the virial is

$$\frac{1}{2}\frac{d^2I}{dt^2} = 2T - \Omega, \tag{1}$$

where T is the kinetic energy of the particles, $-\Omega$ ($\Omega > 0$) the gravitational energy reckoned as zero when all the particles are dispersed at infinity, and $I = \Sigma m r^2$, r being the distance of a typical particle (of mass m) from the centre of symmetry. We may call I the moment of inertia about the centre.

Let $\boldsymbol{V}$ be the mean vector velocity of the particles in any neighbourhood at distance r from the centre, $\boldsymbol{v}$ the random component of velocity of a typical particle

superposed thereon. Then

$$T=\tfrac{1}{2}\Sigma m(\boldsymbol{V}+\boldsymbol{v})^2=\tfrac{1}{2}\Sigma m\boldsymbol{V}^2+\tfrac{1}{2}\Sigma m\boldsymbol{v}^2$$
$$=K+\tfrac{1}{2}\Sigma m\boldsymbol{v}^2,$$

say, where K is the kinetic energy of the mass-motion. Then the equation of the virial becomes

$$\frac{1}{2}\frac{d^2I}{dt^2}=2K+\Sigma m\boldsymbol{v}^2-\Omega. \tag{2}$$

Now if H is the heat energy of the system,

$$H=\int c_v\theta\, dN,$$

where c_v is the specific heat at constant volume, per molecule, θ is the absolute temperature and the integration extends over all the particles counted as molecules. Assume now that the system of particles constitutes an ideal gas. Then if R is the gas-constant, by a familiar result of thermodynamics,

$$c_p-c_v=R,$$

c_p being the specific heat at constant pressure per molecule. Writing, as usual, $c_p/c_v=\gamma$, we have

$$c_v=R/(\gamma-1),$$

whence

$$H=\frac{1}{\gamma-1}\int R\theta\, dN.$$

But by the theorem of equipartition,

$$\tfrac{1}{2}m\boldsymbol{v}^2=\tfrac{3}{2}R\theta.$$

Hence

$$\Sigma m\boldsymbol{v}^2=3\Sigma R\theta=3\int R\theta\, dN=3(\gamma-1)H.$$

Hence the equation of the virial, (2), becomes

$$\frac{1}{2}\frac{d^2I}{dt^2}=2K+3(\gamma-1)H-\Omega. \tag{3}$$

But if E is the total energy of the star,

$$E=K+H-\Omega,$$

when for consistency's sake we have neglected the imprisoned radiant energy.* Eliminating H we get

$$\frac{1}{2}\frac{d^2I}{dt^2}=(3\gamma-4)\Omega-(3\gamma-5)K+3(\gamma-1)E. \tag{4}$$

As a check, in the state of equilibrium $K=0$ and the left-hand side vanishes, so we get

$$E=-\frac{3\gamma-4}{3(\gamma-1)}\Omega,$$

the well-known expression for the total energy of a star (**8**).

* This is taken into account later. See Section 10, p. 524.

3. *Deduction of the luminosity variation.*—In the steady state associated with a given state of pulsation, we have

$$I = \text{const.}, \qquad \Omega = \Omega_0, \qquad K = 0, \qquad E = E_0,$$

and so

$$0 = (3\gamma - 4)\Omega_0 + 3(\gamma - 1)E_0.$$

Put

$$\Omega - \Omega_0 = \delta\Omega, \qquad E - E_0 = \delta E.$$

Then subtracting the last equation from the virial equation in the form (4), we get

$$\tfrac{1}{2}\frac{d^2I}{dt^2} = (3\gamma - 4)\delta\Omega - (3\gamma - 5)K + 3(\gamma - 1)\delta E. \tag{5}$$

Now consider the output of energy from the star's surface. Let L_0 be the luminosity of the associated steady-state configuration, $\mathscr{E}_0$ the rate of generation of sub-atomic energy. Then

$$L_0 = -\frac{dE_0}{dt} + \mathscr{E}_0.$$

When the system is pulsating, let $L_0 + \delta L$ be the luminosity at any phase of the oscillation. We shall assume that the rate of generation of sub-atomic energy is unaffected by the pulsation. Then

$$L_0 + \delta L = -\frac{dE}{dt} + \mathscr{E}_0.$$

Subtracting, we get

$$\delta L = -\frac{d(\delta E)}{dt}.$$

Differentiating the equation of the virial in the form (5) with regard to the time and substituting for $d(\delta E)/dt$ from the last equality, we get

$$\delta L = \frac{1}{3(\gamma - 1)}\left[(3\gamma - 4)\frac{d(\delta\Omega)}{dt} - (3\gamma - 5)\frac{dK}{dt} - \tfrac{1}{2}\frac{d^3I}{dt^3}\right]. \tag{6}$$

4. *Simplification of the formula for the luminosity variation.*—We wish to employ (6) to gain insight into the nature of the luminosity variation without going into the tedious process of integrating the equation of motion of the oscillation. For this purpose certain approximations are necessary. The variation in radial distance of any element, δr, increases from zero at the centre to some value δr_0 at the boundary. If therefore we assume that during the oscillation the ratio $\delta r/r$ remains constant from centre to radius (varying of course with the time), we shall get the right type of behaviour of δr. We shall therefore assume as an approximation valid for calculating the various terms in (6), the homology relation

$$\delta r/r = \delta r_0/r_0, \tag{7}$$

where r_0 is the mean external radius. This approximation is made by Ledoux and Pekeris, and again by Ledoux, in deducing the formulae for the period with which we shall be concerned, and in any case it cannot be far from the truth.

Now for any spherically symmetrical system, the negative gravitational potential energy Ω is of the form

$$\Omega = \alpha\frac{GM^2}{r_0}, \tag{8}$$

where G is the constant of gravitation, M the total mass and α a pure number. The coefficient α depends solely on the relative mass-distribution, and its value is unaltered by a homologous transformation. For example, $\alpha=\frac{3}{5}$ for a homogeneous distribution, and $\alpha=3/(5-n)$ for a complete Emden polytrope of index n. Similarly, we can write

$$I=kMr_0^2, \tag{9}$$

where k is a pure number depending only on the relative mass-distribution. For example, $k=\frac{3}{5}$ for a homogeneous distribution, and $k=0{\cdot}11304$ for an Emden polytrope $n=3$ (**6**). Other values, kindly calculated for me by Mr C. Prasad, are $k=0{\cdot}23225$ for the polytrope $n=2$, and $k=0{\cdot}3921$ for the polytrope $n=1$. As increasing n means increasing mass-concentration towards the centre, we see that with *decreasing* mass-concentration, α decreases and k increases. These directions of change are moreover intuitively obvious. The pure number k, like α, is unaltered by a homologous transformation. Lastly we have for a pulsating star

$$K=\tfrac{1}{2}\Sigma m\dot{r}^2.$$

Since (7) gives

$$\dot{r}/r=\dot{r}_0/r_0,$$

we have on our assumption

$$K=(\tfrac{1}{2}\Sigma mr^2)(\dot{r}_0/r_0)^2=\tfrac{1}{2}kM\dot{r}_0^2. \tag{10}$$

We have now in succession

$$\frac{d(\delta\Omega)}{dt}=\frac{d\Omega}{dt}=-\alpha\frac{GM^2}{r_0^2}\dot{r}_0, \tag{11}$$

$$\frac{dK}{dt}=kM\dot{r}_0\ddot{r}_0, \tag{12}$$

and

$$\frac{d^3I}{dt^3}=kM(6\dot{r}_0\ddot{r}_0+2r_0\dddot{r}_0). \tag{13}$$

Substituting in (6) we get

$$\delta L=-\frac{M}{3(\gamma-1)}\left[(3\gamma-4)\alpha\frac{GM}{r_0^2}\dot{r}_0+(3\gamma-5)k\dot{r}_0\ddot{r}_0+k(3\dot{r}_0\ddot{r}_0+r_0\dddot{r}_0)\right]. \tag{14}$$

5. *Application to small oscillations.*—So far we have not assumed the amplitude of the pulsation to be small. If we now do so, as in Ledoux's work, we have that $\dot{r}_0$, $\ddot{r}_0$, $\dddot{r}_0$ are each of the first order of smallness. Neglecting then the second order product $\dot{r}_0\ddot{r}_0$ we have from (14)

$$\delta L=-\frac{M}{3(\gamma-1)}\left[(3\gamma-4)\alpha\frac{GM}{r_0^2}\dot{r}_0+kr_0\dddot{r}_0\right]. \tag{15}$$

Now assume that the excursion δr_0 of the boundary of the star is harmonic in the time, i. e.

$$\delta r_0=A\sin\sigma t.$$

Then

$$\dot{r}_0=+\sigma A\cos\sigma t, \qquad \dddot{r}_0=-\sigma^3A\cos\sigma t=-\sigma^2\dot{r}_0.$$

Relation (15) now gives

$$\delta L=\frac{\dot{r}_0}{3(\gamma-1)}\left[\sigma^2kMr_0-(3\gamma-4)\alpha\frac{GM^2}{r_0^2}\right], \tag{16}$$

which may be put in the form

$$\delta L = \frac{\dot{r}_0}{3(\gamma-1)r_0}[\sigma^2 I_0 - (3\gamma-4)\Omega_0]. \tag{17}$$

6. *Application of the period formula of Ledoux and Pekeris.*—Formula (17) exhibits the variation in luminosity δL as in phase with the outward surface velocity $\dot{r}_0$ provided the square bracket is positive. It is clear that a proportionality to $\dot{r}_0$ would also have resulted if we had not made the simplifying assumption of a homologous oscillation. It will be noticed that we have nowhere assumed that the oscillations are adiabatic. Proportionality (17) of δL to $\dot{r}_0$ is the principal result of the present paper.

According to the investigations of Ledoux and Pekeris (**6**) (see also (**11**), p. 41), on the hypothesis of adiabatic oscillations of homologous character, the period $2\pi/\sigma$ is given by the formula

$$\sigma^2 = (3\gamma-4)\frac{\Omega_0}{I_0}. \tag{18}$$

It follows that on the adiabatic hypothesis, to the first order,

$$\delta L = 0, \tag{19}$$

for the square bracket in (17) then vanishes.

Thus on the adiabatic hypothesis, there is no fluctuation in the luminosity. The reason is, as shown above, that on the adiabatic hypothesis each volume-element of the system behaves as if adiabatically sealed off from its neighbours, there can be no interchange of radiant energy, and the flow of radiation which gives rise to the luminosity is zero. It should be noted that Ledoux and Pekeris, in deducing (18), explicitly made the approximation of a homologous pulsation, as we have done.

As an example, consider the homogeneous model. The vibration period for radial pulsations, on the adiabatic hypothesis, was found by Ritter ((**10**), and (**11**) p. 5) to be given by

$$\sigma^2 = (3\gamma-4)\frac{g_0}{r_0} = (3\gamma-4)\frac{4\pi}{3}G\rho, \tag{20}$$

whilst for the homogeneous model $k=\frac{3}{5}$ and $\alpha=\frac{3}{5}$. Now formula (16) can be written in the form

$$\delta L = \frac{Mr_0\dot{r}_0}{3(\gamma-1)}\left[k\sigma^2 - (3\gamma-4)\frac{4\pi}{3}G\bar{\rho}\alpha\right], \tag{21}$$

where $\bar{\rho}$ is the mean density $M/\frac{4}{3}\pi r_0^3$. Thus for the homogeneous model δL vanishes identically on the adiabatic hypothesis, in agreement with the general theory.

7. Relation (21) can be put in yet another form. It is well known that for any given mass-distribution, the period P or $2\pi/\sigma$ is inversely proportional to the square root of the mean density, and in fact that

$$P = 2\pi\left[\frac{a}{(3\gamma-4)\pi G\bar{\rho}}\right]^{\frac{1}{2}}, \tag{22}$$

where a is a pure number (a constant) depending on the density-distribution. This constancy of the product $P\rho^{-\frac{1}{2}}$ has been confirmed by Payne-Gaposchkin

and Gaposchkin ((**9**) and (**11**), p. 42) for different types of stars, the constant being greater for the long-period variables than for the Cepheids. Substituting for $P\ (=2\pi/\sigma)$ from (22) in (21) we get

$$\delta L = \frac{(3\gamma-4)\pi G M r_0 \bar{\rho} \dot{r}_0}{3(\gamma-1)} \left[\frac{k}{a} - \tfrac{4}{3}\alpha\right]. \tag{23}$$

Since on the adiabatic hypothesis δL must vanish, formula (23) gives the following relation between the three pure numbers k, α, a for adiabatic pulsations,

$$k = \tfrac{4}{3}\alpha a. \tag{24}$$

This is of course Ledoux and Pekeris's relation (18) in another form. We have seen that (24) is exactly satisfied for the homogeneous model. As a further test, let us apply it to the standard model, the Emden polytrope $n=3$. Here $\alpha = \frac{3}{2}$, $k = 0{\cdot}113$ and so the predicted value of a is 0·0565. Eddington's well-known semi-empirical relation for the adiabatic pulsation period of an Emden polytrope $n=3$ is

$$P = 2\pi \left[\frac{1}{\frac{3}{10}\pi G(3\gamma-4)\rho_c}\right]^{\frac{1}{2}} = 2\pi \left[\frac{1}{\frac{3}{10}\pi G(3\gamma-4)\bar{\rho} \times 54{\cdot}4}\right]^{\frac{1}{2}},$$

which gives the numerical value of a as

$$a = \frac{10/3}{54{\cdot}4} = 0{\cdot}0612.$$

Considering that the adiabatic oscillations of the standard model will not be exactly homologous, the agreement is very fair.

8. *The luminosity-velocity phase relation.*—Since fluctuations of luminosity are observed in nature, we cannot have $\delta L = 0$. It follows that actual pulsations in nature cannot be adiabatic, and since for Cepheids δL is in phase with $\dot{r}_0$, the square bracket in (17) or (23) must be positive for Cepheids. Hence for Cepheids we must have

$$\sigma^2 > (3\gamma-4)\frac{\Omega_0}{I_0} \qquad \text{or} \qquad a < \frac{k}{\frac{4}{3}\alpha}. \tag{25}$$

It is clear, however, from the empirical success of Ledoux and Pekeris's formula (18) for the period that the value of the square bracket in (23) must be numerically very small. In other words, the inequalities (25) must be very nearly equalities. Owing to the uncertainty in the actual density distribution in Cepheids, there is small hope of testing inequalities (25) numerically with any exactitude.

We can see, however, in a general way that the *sign* of the square bracket in (23) will be likely to be reversed (become negative) as we pass to long-period variables, provided these stars have a stronger density-concentration to the centre than Cepheids. For $P\rho^{-\frac{1}{2}}$ according to the Gaposchkin diagram ((**11**), p. 43) has a larger value for long-period variables than for Cepheids, i. e. a is larger for long-period variables; and on the other hand, with increasing mass-concentration α increases and k decreases. The square bracket will, on the other hand, have the best prospect of being positive when k tends to be bigger, α smaller and a smaller. This suggests that Cepheid variables are more nearly homogeneous than long-period variables, in agreement with a conclusion of Eddington's (**3**), arrived at on other grounds.

In numerical confirmation of this we note that the first of the inequalities (25) is certainly satisfied for δ Cephei if it is assumed to be homogeneous with $\gamma = \frac{5}{3}$.

For, taking the data tabulated by Rosseland ((**11**), Table 1, p. 12), with $P = 5{\cdot}37$ days, we have

$$\sigma^2 = \left(\frac{2\pi}{P}\right)^2 = \left[\frac{2\pi}{5{\cdot}37 \times 24 \times 60 \times 60}\right]^2 = 1{\cdot}84 \times 10^{-10},$$

and with $M = 9M_\odot$, $r_0 = 23 \times 10^6$ km., for the homogeneous model

$$(3\gamma - 4)\frac{\Omega_0}{I_0} = (3\gamma - 4)\frac{GM}{r_0^3} = \frac{6{\cdot}66 \times 10^{-8} \times 9 \times 2 \times 10^{33}}{(23 \times 10^{11})^3} = 1 \times 10^{-10}.$$

The general conclusion from the observed phase-relation in Cepheid and long-period variables is that Cepheids approach a homogeneous density-distribution whilst long-period variables are more centrally condensed.

9. *Comparison with observed luminosity-amplitude.*—A further numerical test of the present theory can also be made by noting that according to (23), since the square bracket has been seen to be a very small pure number, we ought to have

$$\frac{\delta L}{M} \ll \frac{3\gamma - 4}{3(\gamma - 1)} \pi G \bar{\rho} r_0 \dot{r}_0. \tag{26}$$

For δ Cephei, Rosseland's tables ((**11**), pp. 12 and 149, Tables 1 and 10) give

$$r_0 = 23 \times 10^6 \text{ km.}, \qquad \dot{r}_0 = 39 \text{ km. sec.}^{-1},$$

whilst for a mass $M = 9M_\odot$ the mean density comes out as

$$\bar{\rho} = 3{\cdot}5 \times 10^{-4} \text{ g. cm.}^{-3}.$$

Hence by (26) we ought to have, with $\gamma = \frac{5}{3}$,

$$\delta L/M \ll 3{\cdot}3 \times 10^8 \text{ ergs sec.}^{-1} \text{ g.}^{-1}.$$

Now the observed amplitude of δL is 1·23 magnitudes. This is a ratio-increase in brightness of $10^{0{\cdot}492}$ or 3·105 times. But according to Eddington's table (**4**), the absolute bolometric magnitude of δ Cephei is $-2{\cdot}17$, and (the Sun's absolute bolometric magnitude being $+4{\cdot}85$) the average energy-emission from δ Cephei is $10^{0{\cdot}4 \times 7{\cdot}02}$, or $10^{2{\cdot}808}$, or 643 times that of the Sun, $3{\cdot}78 \times 10^{33}$ ergs sec.$^{-1}$. Hence the amplitude of the energy-output in δ Cephei is

$$3{\cdot}105 \times 643 \times 3{\cdot}78 \times 10^{33} = 7{\cdot}5 \times 10^{36} \text{ ergs sec.}^{-1},$$

whence

$$\left(\frac{\delta L}{M}\right)_{\text{obs.}} = \frac{7{\cdot}5 \times 10^{36}}{9 \times 2 \times 10^{33}} = 0{\cdot}42 \times 10^3 \text{ ergs sec.}^{-1} \text{ g.}^{-1}.$$

This is indeed small compared with the value of the right-hand side of (26), though one would not have expected the latter to be so very much too large.

The inference from this order-of-magnitude calculation is that the pulsations in δ Cephei must be *very nearly* adiabatic, in agreement with the result based on the discussion of the period. But it is also to be inferred that an adiabatic approximation is quite impotent to ascertain the amplitude of the luminosity-variation theoretically.

10. *Effect of radiation pressure and radiation energy.*—It may be objected to the foregoing that to leave out terms due to radiation in a discussion of the changes in radiant energy emitted by a star is to act *Hamlet* without the Prince of Denmark. I therefore proceed to give the necessary modifications in the formulae caused by taking account of radiation.

The mechanical equation of motion,

$$\rho \frac{d^2r}{dt^2} = -\frac{\partial p}{\partial r} - \frac{\partial p'}{\partial r} - \frac{GM(r)}{r^2}\rho,$$

when multiplied by $4\pi r^3\,dr$ and integrated over the whole volume of the star, gives after an integration by parts

$$\int r \frac{d^2r}{dt^2} dM(r) = 3\int (p+p')\,dv - \int \frac{GM(r)\,dM(r)}{r},$$

or, carrying out the usual virial transformations on the left-hand side,

$$\frac{1}{2}\frac{d^2I}{dt^2} = 2K + 3\int P\,dv - \Omega, \qquad (27)^*$$

where K is the kinetic energy of the mass-motion only, and p, p', P are gas-pressure, radiation pressure and total pressure, respectively.

Now for gaseous material,

$$\int p\,dv = \int (R/\mu)\rho\theta\,dv = \int R\theta\,dN = (\gamma-1)H,$$

as in Section 2. We may now put, in Eddington's notation,

$$p = \beta P, \qquad p' = (1-\beta)P,$$

and then

$$3\int P\,dv = \frac{3(\gamma-1)}{\beta}H.$$

Hence the equation of the virial now is

$$\frac{1}{2}\frac{d^2I}{dt^2} = 2K - \Omega + \frac{3(\gamma-1)}{\beta}H. \qquad (28)$$

In this we have assumed β constant; if it is not constant, it will be sufficient to give it a mean value.

The total energy of the star is now given by

$$E = K - \Omega + H + U, \qquad (29)$$

where U, the total radiant energy, is given by

$$U = 3\int p'\,dv = 3\frac{1-\beta}{\beta}(\gamma-1)H.$$

Eliminating H between the equation of the virial (28), and the equation of energy (29), we get

$$\frac{1}{2}\frac{d^2I}{dt^2} = 2K - \Omega + \frac{3(\gamma-1)}{\beta}\left[E - K + \Omega\right] \Big/ \left[1 + 3\frac{1-\beta}{\beta}(\gamma-1)\right],$$

or

$$\frac{1}{2}\frac{d^2I}{dt^2} = \frac{3(\gamma-1)E + (3\gamma-4)\beta\Omega + [\beta(5-3\gamma) + 3(1-\beta)(\gamma-1)]K}{\beta + 3(\gamma-1)(1-\beta)}. \qquad (30)$$

* There should be no confusion between this use of P to denote total pressure and the earlier use of P to denote the period.

By the same procedure as before, we now get a luminosity formula of the form

$$\delta L = \frac{1}{3(\gamma-1)}\left[(3\gamma-4)\beta\frac{d(\delta\Omega)}{dt} - \tfrac{1}{2}\{\beta+3(\gamma-1)(1-\beta)\}\frac{d^3I}{dt^3} + \{\beta(5-3\gamma)+3(1-\beta)(\gamma-1)\}\frac{dK}{dt}\right]. \quad (31)$$

If we now neglect second-order terms, this yields to the same degree of approximation as before,

$$\delta L = \frac{\dot{r}_0}{3(\gamma-1)}\left[\{\beta+3(\gamma-1)(1-\beta)\}\sigma^2 kMr_0 - (3\gamma-4)\beta\alpha\frac{GM^3}{r_0^2}\right]$$
$$= \frac{\beta\dot{r}_0}{3(\gamma-1)r_0}\left[\left\{1+3(\gamma-1)\frac{1-\beta}{\beta}\right\}\sigma^2 I_0 - (3\gamma-4)\Omega_0\right]. \quad (32)$$

For *adiabatic* pulsations, for which $\delta L = 0$, we get the approximate formula

$$\left(\frac{2\pi}{P}\right)^2 = \sigma^2 = \frac{3\gamma-4}{1+3(\gamma-1)(1-\beta)/\beta}\frac{\Omega_0}{I_0}. \quad (33)$$

The existence of appreciable radiation pressure accordingly gives a smaller value of σ, or a lengthened period, for purely adiabatic oscillations; e. g. for $1-\beta = 0{\cdot}45$ (Eddington's original value for δ Cephei), the effect of radiation pressure on adiabatic oscillations is to lengthen the period in the ratio

$$\left[1+3(\tfrac{5}{3}-1)\frac{0{\cdot}45}{0{\cdot}55}\right]^{\frac{1}{2}} = 1{\cdot}28.$$

The actual oscillation of a variable star cannot however, as we have seen, be adiabatic, and thus the square bracket in (32) cannot vanish for an actual pulsating variable. We now see that the effect of the presence of radiation pressure is to increase the coefficient of $\sigma^2 I_0$, and so stars with high radiation pressure are more likely to give δL in phase with $\dot{r}_0$ than stars with low radiation pressure.

Thus Cepheids are theoretically likely to give δL in phase with $\dot{r}_0$, whilst long-period variables, with relatively low radiation pressure, are likely to give δL in phase with $-\dot{r}_0$, as observed.

11. *Effect of rotation.*—Following Ledoux, we can apply our luminosity formula to give the effect of rotation on the period of adiabatic pulsations. When rotation is present, the macroscopic kinetic energy K is the sum of two terms, the kinetic energy of the pulsational motion and the kinetic energy of the rotational motion. Thus we may write

$$K = \tfrac{1}{2}kM\dot{r}_0^2 + \tfrac{1}{2}\kappa Mr_0^2\omega^2,$$

where κr_0^2 is the square of the radius of gyration about the axis of rotation. We have then

$$\frac{dK}{dt} = kM\dot{r}_0\ddot{r}_0 + \kappa M(r_0\dot{r}_0\omega^2 + r_0^2\omega\dot{\omega}),$$

where $\dot{\omega}$ is determined from the principle of constancy of angular momentum in the form

$$0 = \frac{d}{dt}[\kappa Mr_0^2\omega] = \kappa M[2r_0\dot{r}_0\omega + r_0^2\dot{\omega}].$$

Eliminating $\dot{\omega}$, we have

$$\frac{dK}{dt} = kM\dot{r}_0\ddot{r}_0 + \kappa M(r_0\dot{r}_0\omega^2 - 2r_0\dot{r}_0\omega^2)$$
$$= kM\dot{r}_0\ddot{r}_0 - \kappa M r_0\dot{r}_0\omega^2.$$

The first term here is of the second order of small quantities. Neglecting it, we have

$$\frac{dK}{dt} = -\frac{J\omega}{r_0}\dot{r}_0,$$

where J is the angular momentum, $\kappa M r_0^2\omega$. Inserting this term in (6) we have

$$\delta L = \frac{1}{3(\gamma-1)}\left[(3\gamma-4)\frac{d(\delta\Omega)}{dt} + (3\gamma-5)\dot{r}_0\frac{J\omega}{r_0} - \frac{1}{2}\frac{d^3I}{dt^3}\right],$$

whence, as in (16),

$$\delta L = \frac{\dot{r}_0}{3(\gamma-1)}\left[\sigma^2 kMr_0 - (3\gamma-4)\alpha\frac{GM^2}{r_0^2} + (3\gamma-5)\frac{J\omega}{r_0}\right]$$
$$= \frac{\dot{r}_0}{3(\gamma-1)r_0}[\sigma^2 I_0 - (3\gamma-4)\Omega_0 + (3\gamma-5)J\omega]. \qquad (34)$$

It follows that on the adiabatic hypothesis, for which $\delta L = 0$, the period $2\pi/\sigma$ is given in the presence of rotation by the formula

$$\sigma^2 = \frac{1}{I_0}[(3\gamma-4)\Omega_0 + (5-3\gamma)J\omega]. \qquad (35)$$

This is the formula found by Ledoux ((**5**), p. 152) for adiabatic oscillations and confirmed by Cowling and Newing (**1**). Our formula (34) for the luminosity variation shows that for $\gamma < \frac{5}{3}$, the effect of the rotation term is to diminish the luminosity variation in the case of Cepheids and to increase it in the case of long-period variables. When radiation pressure is appreciable, (32) modifies by incorporation of the rotation term to

$$\delta L = \frac{\beta\dot{r}_0}{3(\gamma-1)r_0}\left[\left\{1 + 3(\gamma-1)\frac{1-\beta}{\beta}\right\}\sigma^2 I_0 - (3\gamma-4)\Omega_0 - \left\{(5-3\gamma) + 3(\gamma-1)\frac{1-\beta}{\beta}\right\}J\omega\right]. \qquad (36)$$

University Observatory,
Oxford:
1949 *June* 2.

References

(1) T. G. Cowling and R. A. Newing, *Ap. J.*, **109**, 154, 1949.
(2) A. S. Eddington, *M.N.*, **79**, 2 and 177, 1918.
(3) A. S. Eddington, *The Observatory*, **64**, 233, 1942.
(4) A. S. Eddington, *Internal Constitution of the Stars*, p. 182, Table 24, 1926.
(5) P. Ledoux, *Ap. J.*, **102**, 143, 1945.
(6) P. Ledoux and C. L. Pekeris, *Ap. J.*, **94**, 124, 1941.
(7) E. A. Milne, *Nature*, **163**, 703, 1949.
(8) E. A. Milne, *Handb. d. Ap.*, Bd. III/1, p. 199, 1930.
(9) C. H. Payne-Gaposchkin and S. Gaposchkin, *Proc. Amer. Phil. Soc.*, **86**, 329, 1942.
(10) A. Ritter, *Wied Ann.*, **8**, 172, 1879.
(11) S. Rosseland, *The Pulsation Theory of Variable Stars*, Clarendon Press, 1949.

ON THE ANHARMONIC PULSATIONS OF THE STANDARD MODEL

Chandrika Prasad

(Communicated by E. A. Milne)

(Received 1949 July 6)

Summary

A method of solving the equations of anharmonic pulsation by the use of Fourier series is developed. This is applied to the standard model. It is found that the first two overtones give a considerable skewness to the radial velocity curve, and the effect of the totality of overtones is discussed. It is also found that the increase in the period is slight.

---

1. *Introduction.*—The theory of anharmonic pulsations for the Cepheid variables was developed by Rosseland * to explain the skewness of the radial velocity curve, and also the longer period of oscillation observed than predicted by the standard model. Working on the homogeneous model of the star and taking the case of a single excited mode, he was able to show that on retaining the second-order term in the equation of motion, the period lengthened and the velocity curve acquired a skewness. But these results were only qualitative, being quite insufficient to account for the large lengthening and skewness observed in Cepheids for the comparatively small amplitude of pulsation.

Rosseland expected that better agreement would be obtained by retaining higher-order terms. But his expectations were belied by the work of Bhatnagar and Kothari † who completely integrated the equation of a single excited mode for the homogeneous model for $\gamma = \frac{5}{3}$. Similar conclusions were reached by Roy ‡ working on more general lines. His work also indicated that inclusion of the second mode of oscillation would bring us nearer the desired result. The complete solution of the equations for the fundamental mode and the first overtone were not attempted for a considerable time, probably due to the arduous numerical integration of the equations involved. But recently a solution has been given for the standard model by Schwarzschild and Savedoff.§ Their results show that the lengthening of the period is very small for the observed pulsation-amplitude of 0·06; and though the skewness obtained is not small, it is yet insufficient to account for the observed skewness.

In the present paper a method of solving the equations of anharmonic pulsation by the use of Fourier series is developed. The method is then applied to the equations for the higher overtones of the standard model. The use of Fourier series has been indicated by Rosseland || as a means of solving the approximate equation for an overtone. But he apparently did not realize that they can be

* S. Rosseland, *M.N.*, **103**, 239, 1943.
† P. L. Bhatnagar and D. S. Kothari, *M.N.*, **104**, 292, 1944.
‡ S. K. Roy, *Bull. Calcutta Math. Soc.*, **36**, 109, 1944.
§ M. Schwarzschild and M. P. Savedoff, *Ap. J.*, **109**, 208, 1949.
|| S. Rosseland, *The Pulsation Theory of Variable Stars*, Oxford, p. 106, 1949.

used to solve the equations in their complete generality; nor did he give a method for evaluating the coefficients.

The results show that the effects of the overtones are considerable. The first two overtones give a skewness of the velocity curve in fair agreement with observation. But there is an indication that the skewness may even exceed the observed value in some cases. Also the lengthening of the period is slight. On both these accounts a more homogeneous model for the Cepheid stars is indicated. A similar conclusion has recently been reached by Milne* on different grounds.

2. *Equations of the problem.*—The exact equation of motion for radial adiabatic oscillations of a non-rotating star is†

$$\rho_0 r_0 \ddot{r}_1 = -(1+r_1)^2 \frac{\partial}{\partial r_0}[P_0(1+r_1)^{-2\gamma}(1+r_1+r_0 r_1')^{-\gamma}] + P_0'(1+r_1)^{-2}, \qquad (1)$$

where P_0, ρ_0 and r_0 are the equilibrium values of pressure, density and distance from the centre, γ is the ratio of specific heats, and r_1 is the relative displacement defined by $r = r_0(1+r_1)$. The dashes denote differentiation with respect to r_0. The relative displacement may be written as‡

$$r_1 = \sum_k \eta_k(r_0) q_k(t), \qquad (2)$$

where η's are the solutions, already determined, of the equation of motion for infinitesimal amplitude, for the various modes of oscillation, and q's are functions of time which are to be determined. The q's reduce to sine and cosine functions for infinitesimal amplitude. We shall take the η's to be normalized to unity at the surface, so that Σq_k is the relative displacement at the star's surface.

Multiply equation (1) by $r_0^3 \eta_k$ and integrate with respect to r_0 from the centre to the boundary R of the star. Due to the orthogonality of the η-functions, linear cross-terms drop out and we obtain on expanding the right-hand side up to second order, and performing some partial integrations,

$$\ddot{q}_k + \sigma_k^2 q_k = \frac{1}{I_k}\Big[-\tfrac{1}{2}(3\gamma-4)(3\gamma+1)\int_0^R P_0' r_0^3 \eta_k r_1^2 dr_0 + \tfrac{1}{2}\gamma(3\gamma-1)\int_0^R P_0 r_0^4 \eta_k r_1'^2 dr_0$$
$$+\gamma(3\gamma-1)\int_0^R P_0 r_0^4 \eta_k' r_1 r_1' dr_0 + \tfrac{1}{2}\gamma(\gamma+1)\int_0^R P_0 r_0^5 \eta_k' r_1'^2 dr_0 \Big], \qquad (3)$$

where $2\pi/\sigma_k$ is the period of oscillation for infinitesimal amplitude corresponding to the mode η_k, and

$$I_k = \int_0^R \rho_0 r_0^4 \eta_k^2 dr_0. \qquad (4)$$

The value of r_1 from (2) is to be substituted in the integrals on the right-hand side of (3), and they are to be arranged in various products of the q's. Further, it is found convenient to normalize the time by taking

$$t = \tau/\sigma_1, \qquad (5)$$

so that the period of the fundamental mode becomes 2π for infinitesimal amplitude.

* E. A. Milne, *M.N.*, **109**, 517, 1949.
† S. Rosseland, *The Pulsation Theory of Variable Stars*, Oxford, p. 97, 1949.
‡ *Ibid.*, p. 53.

Then (3) may be written as

$$\frac{d^2q_k}{d\tau^2}+\beta_k q_k=\frac{\gamma}{\sigma_1^2 I_k}[\Sigma D_{iik}q_i^2+2\Sigma D_{ijk}q_iq_j], \tag{6}$$

where

$$\beta_k=\sigma_k^2/\sigma_1^2, \tag{7}$$

$$\begin{aligned}D_{ijk}=&-\tfrac{1}{2}\left(3-\frac{4}{\gamma}\right)(3\gamma+1)\int_0^R P_0'r_0^3\eta_i\eta_j\eta_k\,dr_0\\&+\tfrac{1}{2}(3\gamma-1)\int_0^R P_0r_0^4(\eta_i\eta_j'\eta_k'+\eta_i'\eta_j\eta_k'+\eta_i'\eta_j'\eta_k)\,dr_0\\&+\tfrac{1}{2}(\gamma+1)\int_0^R P_0r_0^5\eta_i'\eta_j'\eta_k'\,dr_0,\end{aligned} \tag{8}$$

and i, j, k take integral values from 1 onwards. The D's are constants depending upon the particular stellar model chosen. Only the terms involving q_1 will be the terms of importance on the right-hand side of (6); the rest would be higher-order terms. This is seen below.

3. *Method of solution.*—To illustrate the method of solution we shall take only the equations for q_1 and q_2. These may be written as

$$\frac{d^2q_1}{d\tau^2}+q_1=A_1q_1^2+2B_1q_1q_2+C_1q_2^2, \tag{9}$$

$$\frac{d^2q_2}{d\tau^2}+\beta_2q_2=A_2q_1^2+2B_2q_1q_2+C_2q_2^2, \tag{10}$$

where $A_1, B_1, \ldots$ are known constants.

We require solutions of these equations which are periodic with the same period for q_1 and q_2. Hence we put

$$q_1=a_0+a_1\cos n\tau+a_2\cos 2n\tau+a_3\cos 3n\tau+\ldots, \tag{11}$$

$$q_2=b_0+b_1\cos n\tau+b_2\cos 2n\tau+b_3\cos 3n\tau+\ldots, \tag{12}$$

where $a_0, a_1, \ldots, b_0, b_1, \ldots$ and n are constants to be determined. Sine terms are excluded as by a choice of origin the solution can be made symmetrical in time.*

These values of q_1 and q_2 are substituted in equations (9) and (10) and the products of cosines are expressed as sums of cosines. Then putting the constant term and the coefficients of $\cos kn\tau$ for different k, separately equal to zero, we get the following equations from (9):—

$$\begin{aligned}-a_0+A_1[a_0^2+\tfrac{1}{2}a_1^2+\tfrac{1}{2}a_2^2+\ldots]+2B_1[a_0b_0+\tfrac{1}{2}a_1b_1+\tfrac{1}{2}a_2b_2+\ldots]\\+C_1[b_0^2+\tfrac{1}{2}b_1^2+\tfrac{1}{2}b_2^2+\ldots]=0,\end{aligned} \tag{13}$$

$$\begin{aligned}(n^2-1)a_1+A_1[2a_0a_1+a_1a_2+a_2a_3+\ldots]+2B_1[a_1b_0+a_0b_1+\tfrac{1}{2}(a_1b_2+a_2b_1)\\+\tfrac{1}{2}(a_2b_3+a_3b_2)+\ldots]+C_1[2b_0b_1+b_1b_2+b_2b_3+\ldots]=0,\end{aligned} \tag{14}$$

$$\begin{aligned}(k^2n^2-1)a_k+A_1\left[\tfrac{1}{2}\sum_{i=0}^{k}a_ia_{k-i}+\sum_{i=0}^{\infty}a_ia_{k+i}\right]\\+B_1\left[\sum_{i=0}^{k}a_ib_{k-i}+\sum_{i=0}^{\infty}(a_ib_{k+i}+a_{k+i}b_i)\right]\\+C_1\left[\tfrac{1}{2}\sum_{i=0}^{k}b_ib_{k-i}+\sum_{i=0}^{\infty}b_ib_{k+i}\right]=0,\end{aligned} \tag{15}$$

* Schwarzschild and Savedoff, *loc. cit.* They show that the solution must be unique. Thus no other solution (e. g. one asymmetrical in time) is possible which is periodic with the same period for q_1 and q_2; barring, of course, a higher mode.

and similar equations from (10), which are the same as above but with $(k^2n^2-\beta_2)b_k$ replacing $(k^2n^2-1)a_k$ and A_2, B_2, C_2 replacing A_1, B_1, C_1 respectively. These equations give a_0, b_0, ... in terms of each other and n. At a first glance these equations appear formidable, but in reality they are not so, as the terms fall off rapidly.

To solve these equations we shall suppose that a_1 is a known small quantity and we shall determine the other coefficients in terms of a_1. Considering equation (13) we see that a_0 is a small quantity of the second order, for all the other terms are squares or products, and they contain amongst them the term a_1^2. Similarly we find that a_2 is of the second order, a_3 is of the third order, and in general $a_k(k>1)$ is of the kth order. Equation (14) then shows that (n^2-1) is a small quantity of the second order. Considering the equations for b's, we find that b_0 is of the second order, b_1 of the third order, b_2 of the second order, and in general $b_k(k>1)$ of the kth order.

A first approximation to the solution of equations (13)–(15) is easy to obtain, and we find that

$$\left.\begin{aligned} a_0&=\tfrac{1}{2}A_1a_1^2, & a_2&=-\tfrac{1}{6}A_1a_1^2, & a_3&=\tfrac{1}{16}\left(\tfrac{1}{3}A_1^2+\frac{B_1A_2}{4-\beta_2}\right)a_1^3,\\ b_0&=\tfrac{1}{2}\left(\frac{A_2}{\beta_2}\right)a_1^2, & b_1&=\left\{\tfrac{5}{6}A_1+\frac{(8-3\beta_2)B_2}{2\beta_2(4-\beta_2)}\right\}\left(\frac{A_2}{\beta_2-1}\right)a_1^3,\\ b_2&=-\tfrac{1}{2}\left(\frac{A_2}{4-\beta_2}\right)a_1^2, & b_3&=\tfrac{1}{2}\left(\tfrac{1}{3}A_1+\frac{B_2}{4-\beta_2}\right)\left(\frac{A_2}{9-\beta_2}\right)a_1^3, \end{aligned}\right\} \qquad (16)$$

and

$$n^2=1-\left\{\tfrac{5}{6}A_1^2+\frac{(8-3\beta_2)B_1A_2}{2\beta_2(4-\beta_2)}\right\}a_1^2,$$

correct up to third order. By substituting these values again in equations (13)–(15) we can obtain a second approximation correct up to fifth order.

A consideration of the equations (16) leads to useful conclusions. It is seen that the inclusion of the cross-terms q_1q_2 in the equations does not affect the second-order terms but affects the third-order terms. A substitution of the numerical values for the standard model shows that the effect of the term in q_1q_2 upon q_1 is negligible, but q_2 is considerably modified by it. The terms in q_2^2 do not affect the solution up to the third order and the fourth-order correction due to them is small as the coefficients C_1 and C_2 are comparatively smaller. Application of (16) also shows that for small amplitudes the skewness of the velocity curve is a small quantity of the first order, whereas the increase in the period is a quantity of the second order. This increase is not due to the term in q_1^2 alone, but is also due to the term in q_1q_2.

A solution for the standard model of the equations (9) and (10), excluding the terms in q_2^2, is found to be

$$q_1=0{\cdot}001765+0{\cdot}05881\cos n\tau-0{\cdot}000588\cos 2n\tau+0{\cdot}000014\cos 3n\tau, \qquad (17)$$

$$q_2=0{\cdot}004788+0{\cdot}002467\cos n\tau-0{\cdot}004036\cos 2n\tau+0{\cdot}000150\cos 3n\tau, \qquad (18)$$

where

$$n^2=0{\cdot}995237.$$

The coefficients above are calculated up to a second approximation only. A comparative table of these solutions and those found by Schwarzschild

and Savedoff is given below. It is seen that the difference for q_1 is negligible. The difference for q_2 is about half due to the neglect of the term $C_2q_2^2$ in (10), and half due to imperfect approximation.

Comparison of Solutions

τ	Our approximation		Exact solution	
	q_1	q_2	q_1	q_2
0·0	0·06000	0·00337	0·06000	0·00360
0·6	0·05013	0·00532	0·05014	0·00553
1·2	0·02365	0·00851	0·02366	0·00865
1·8	−0·01081	0·00797	−0·01081	0·00802
2·4	−0·04141	0·00277	−0·04140	0·00277
3·0	−0·05697	−0·00165	−0·05695	−0·00167

The above method of solution can easily be extended to a more general case, including higher overtones. In the present paper the word "overtone" has been used for the oscillations $q_2, q_3, \ldots$. They differ from the usual sense in which the word is used in not having their periods one-half, one-third, etc., of the period of the fundamental. Actually, in the above solution q_2 is chosen to have the same period as the fundamental oscillation q_1. But the predominant term in q_2 is $b_2 \cos 2n\tau$, which has half the period of the fundamental. The same is the case for $q_3, q_4, \ldots$; the predominant term there has half the period of q_1.

It may also be pointed out here that the overtones will always be present when the fundamental oscillation q_1 is of finite amplitude. They follow as a consequence of the equations of anharmonic pulsation. We cannot escape them by saying that the other modes are not excited. The case of a secondary excited mode will be a quite different problem, showing two predominant periods.

4. *Effect of the second overtone.*—We shall now apply this method of solution to find the effect of the second overtone on the anharmonic pulsations of the standard model. The coefficients in the differential equation have to be calculated by numerical integration. For this purpose the values of the pressure and density were obtained from *B.A. Mathematical Tables*, Vol. II, while the values of η_1, η_2 and η_3 were taken from Schwarzschild's work on the overtones of the standard model.* Some of the coefficients are already calculated in the paper of Schwarzschild and Savedoff.

The equations for the first three modes, for $\gamma = \frac{5}{3}$, are found to be

$$\ddot{q}_1 + q_1 = 0{\cdot}9897q_1^2 + 0{\cdot}6547q_1q_2 + 0{\cdot}0796q_1q_3, \tag{19.1}$$

$$\ddot{q}_2 + 1{\cdot}8354q_2 = 4{\cdot}738q_1^2 + 7{\cdot}533q_1q_2 + 2{\cdot}435q_1q_3, \tag{19.2}$$

$$\ddot{q}_3 + 3{\cdot}079q_3 = 4{\cdot}301q_1^2 + 18{\cdot}175q_1q_2 + 16{\cdot}855q_1q_3, \tag{19.3}$$

where $\ddot{q}_k$ stands for $d^2q_k/d\tau^2$. Only the second and third-order terms have been retained in the equations. The terms in q_2^2, q_3^2 and q_2q_3 are neglected to save the labour of calculating their nine coefficients. We saw in the preceding section that the error due to neglect of q_2^2 was small. The total error introduced here by the neglect of these terms is estimated to be less than one per cent of the amplitude. Moreover, their inclusion will serve no great purpose unless accompanied by the inclusion of terms in q_1^3, $q_1^2q_2$ and $q_1^2q_3$, which are of the same order.

* M. Schwarzschild, *Ap. J.*, **94**, 245, 1941.

A solution of these equations, chosen so as to make the initial amplitude of the fundamental [i. e. $q_1(0)$] equal to 0·06, is

$$q_1 = 0{\cdot}001764 + 0{\cdot}058810 \cos n\tau - 0{\cdot}000588 \cos 2n\tau + 0{\cdot}000017 \cos 3n\tau - 0{\cdot}000001 \cos 4n\tau, \quad (20.1)$$

$$q_2 = 0{\cdot}004810 + 0{\cdot}002113 \cos n\tau - 0{\cdot}004072 \cos 2n\tau + 0{\cdot}000253 \cos 3n\tau - 0{\cdot}000012 \cos 4n\tau, \quad (20.2)$$

$$q_3 = 0{\cdot}003023 + 0{\cdot}000823 \cos n\tau - 0{\cdot}010205 \cos 2n\tau + 0{\cdot}001256 \cos 3n\tau - 0{\cdot}000065 \cos 4n\tau, \quad (20.3)$$

where

$$n^2 = 0{\cdot}995406. \quad (21)$$

If we denote by q_b the total amplitude due to the first three modes at the surface of the star, then

$$q_b = 0{\cdot}009597 + 0{\cdot}061746 \cos n\tau - 0{\cdot}014865 \cos 2n\tau + 0{\cdot}001526 \cos 3n\tau - 0{\cdot}000078 \cos 4n\tau. \quad (22)$$

The shape of the radial velocity curve obtained from this is shown in the diagram. If we denote by K the ratio of the time of rise of the velocity, from minimum to maximum, to the total period, then K for the above curve is 0·307, which is quite near the observed value, which is 0·20 for δ Cephei and 0·24 for η Aquilae.* But the increase in period is slight, being only 0·23 per cent.

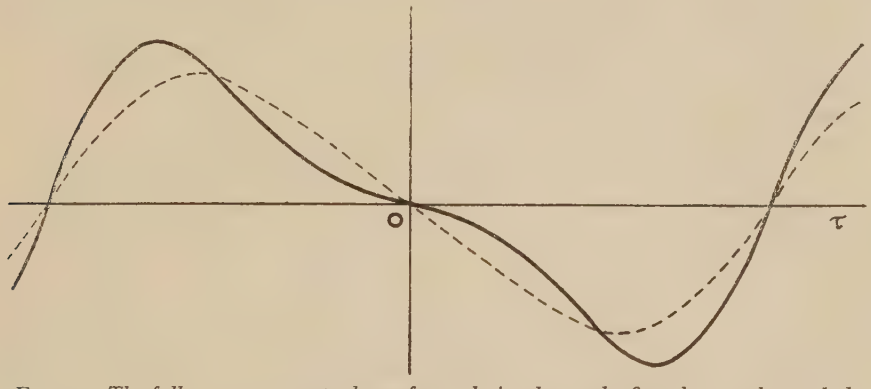

FIG. 1.—*The full curve represents the surface velocity due to the first three modes, and the dotted curve that due to the first two modes.*

5. *Higher overtones.*—We thus see that the effect of the second overtone is considerable; and it leads us to think that the effect of the higher overtones may not be negligible. The equation of the fourth mode of oscillation for the standard model is

$$\ddot{q}_4 + 4{\cdot}697 q_4 = 0{\cdot}0095 q_1^2 + 29{\cdot}24 q_1 q_4 \quad (23)$$

up to two terms. Due to the small coefficient of q_1^2, q_4 is quite small (its amplitude being about one two-hundredth of q_2). Also because

$$\beta_4 = 4{\cdot}697 > 4,$$

the contribution of q_4 is to make the skewness less. That this is an exceptional mode is seen by considering the next mode. The equation for it up to two terms is

$$\ddot{q}_5 + 6{\cdot}67 q_5 = -0{\cdot}049 q_1^2 + 44{\cdot}7 q_1 q_5. \quad (24)$$

It is seen that the coefficient of q_1^2 has numerically increased and is of negative sign. Thus q_5 has a larger amplitude than q_4 (about one-hundredth of q_2) and makes the skewness greater.

* Schwarzschild and Savedoff, *loc. cit.*

Complete solutions for q_4 and q_5 are not given, as they would involve calculation of 12 more coefficients in the equations, before proceeding to solve them; and the contributions of q_4 and q_5 are not large enough to matter much. But the trend of the equations (23) and (24) leaves us in doubt. The coefficient of q_1^2 becomes small in (23), only to turn negative and increase (numerically) in the equation for the higher modes. How far the amplitudes of the modes will go increasing, it is not possible to judge. It may also be noted that the coefficients of $q_1 q_k$ show a continuous increase as we pass to the equations of higher modes. Though it would not be unreasonable to assume that the effect of the modes higher than the third is small, it cannot be definitely stated, without an extensive investigation of higher modes, that the sum of all the modes will be convergent.

The large effect of the second overtone is partly a resonance phenomenon, as β_3 is near the value 4. For smaller values of γ the resonance would be more perfect and consequently q_3 would be larger. This would mean a greater skewness of the radial velocity curve. A rough consideration of the cubic terms, which we have neglected, shows that their contribution would be to make the skewness still greater.

The large coefficients in the equations, and therefore the consequent large amplitudes of the overtones, are mainly due to the tenuous outer parts of the standard model and the steep rise of the overtone amplitudes in these outer parts. Also, the skewness of the velocity curve is a surface effect, the curve becoming rapidly symmetrical as we go inwards. At 90 per cent of the radius the value of K increases to 0·39, and at 83 per cent of the radius to 0·46. For a perfectly symmetrical curve $K = 0{\cdot}5$.

6. *Conclusions.*—We have found that the overtones in a standard model considerably modify the shape of the radial velocity curve, though they have negligible effect on the period. The coefficients in the differential equations for higher overtones show a continuous increase, which renders us unable to reach a definite conclusion about the total effect of the overtones. The skewness obtained by considering the first two overtones is of the order observed, though it appears that it will exceed the observed values if a value of γ lower than $\frac{5}{3}$ is taken (as would be the case in the presence of radiation pressure). On this account, as well as that of the period, a model with less central condensation is indicated.

Two of these are studied in a subsequent paper, where it is found that the overtones are absent in a homogeneous star and the skewness is small. For the model in which the density varies inversely as the square of the distance, the first overtone, which is the predominant one, gives a skewness in the wrong direction. This throws light on the peculiarity of the standard model in having the period of its first overtone (and even the second one for $\gamma = \frac{5}{3}$) greater than half. It would be of interest to consider the anharmonic pulsations of a model, like the polytrope with index 2, which combines a lower density condensation with small density in the outer parts.

I am grateful to Professor E. A. Milne for kindly going through the paper and suggesting several improvements. I also wish to express my thanks to Professor H. H. Plaskett for allowing me the use of the facilities at the Observatory for calculation.

Queen's College,
Oxford:
1949 *July* 2.

ON THE ABERRATIONS OF THE FIELD-FLATTENED SCHMIDT CAMERA

E. H. Linfoot and P. A. Wayman

(Received 1949 August 16)

Summary

The paper develops a general aberration theory of the field-flattened Schmidt camera and applies it to compare the optical performances of the field-flattened Schmidt camera and of the plain Schmidt in a typical special case.

In Section 2 the aberration theory of the plain Schmidt is summarized, and the off-axis coma which results from a given shift of the corrector plate along the axis is evaluated. In Section 3, the aberrations introduced by the field-flattener are calculated, and the aberration-function is derived for the system obtained by adding a field-flattener to a plain Schmidt. In Section 4, the question of balancing the residual aberrations is considered, the amount of plate-shift needed to balance the coma over a given field is determined and it is shown that nothing appreciable is gained, in field-flattened Schmidt cameras of focal ratio $f/3$, by departing from a strictly zero value for the Petzval sum. Finally, in Section 5, comparison is made by means of spot diagrams between the optical performances of an $f/3$ plain Schmidt working over a five-degree field, of the " convertible " system obtained by adding a field-flattener to this and readjusting the corrector plate, and of the non-convertible field-flattened Schmidt of the same field and aperture-ratio in which corrector plate and field-flattener are designed together. It appears that the addition of a field-flattener to an $f/3$ plain Schmidt lowers its performance by unbalancing its colour correction. In the non-convertible field-flattened Schmidt, the aberrations are not much larger than in the plain Schmidt, but they include a small amount of coma.

1. *Introduction.*—As is well known, the field-surface of the ordinary Schmidt camera is not flat, but is convex towards the mirror, its radius of curvature being usually taken as equal to the focal length f. To the Seidel order of approximation, this agrees with the Petzval radius of the system, and the image-surface coincides with the Petzval surface. For the Petzval radius is $f[1+O(\lambda^2)]$ and the distance between the adopted image-surface and the Petzval surface is $O(f\lambda^4)$ all over the field, where $\lambda=1/4F$ and F is the focal ratio.

A thin plano-convex lens placed close in front of the field-surface will change the Petzval sum of the system without introducing appreciable Seidel aberrations into the images. The system remains an anastigmat; that is to say the "deviations" of the points where the rays meet its focal surface are still $O(f\lambda^5)$. Its field-surface therefore continues to coincide with its Petzval surface to the Seidel order of approximation.

If n_1 denote the refractive index of the plano-convex lens and f_1 its focal length, the Petzval contribution of the lens is $-\frac{1}{n_1 f_1}$ and the Petzval curvature of the whole system is $\frac{1}{f}-\frac{1}{n_1 f_1}$. To the order of approximation here in question, the variation of n_1 from one part of the spectrum to another is negligible; more precisely, the effect of this variation is a contribution $O(f\lambda^5)$ to the deviations in the adopted image-surface. By choosing $f_1=f/n_1$, we therefore obtain an

anastigmat with a flat Petzval surface, and consequently a flat image-surface. This system is known as the field-flattened Schmidt camera.

Two different questions now present themselves for investigation. The first, and possibly the more important from the practical astronomer's point of view, is: given a Schmidt camera, how much is necessarily lost in image-quality when a field-flattener is added to the system and how can this loss be minimized?

In Sections 2–4 of the present paper, general formulae are obtained which provide a means of answering this question and in the first part of Section 5 they are applied to discuss the typical special case of an $f/3$ camera working over a 5° field. It appears that the ordinary Petzval equation

$$\frac{1}{f}-\frac{n_1-1}{n_1 r_1}=0$$

provides a sufficiently accurate basis for the design of the field-flattener (r_1 here denoting the radius of curvature of its convex surface) and that, as in the plain Schmidt of this focal ratio and field-size, colour-error dominates the aberrations. There is a considerable drop in performance, due to increased colour-error, when the field-flattener is added. The off-axis coma, when suitably balanced by moving the corrector plate a small distance towards the mirror, is not so large as to affect visibly the symmetry of the photographic star-images formed by the system, but not so small that we can rule out of consideration the possibility of magnitude-errors in their measured positions. An experimental investigation of this question seems desirable.

Field-flattened systems of the above type are convertible, in the sense that they can be used either with or without the field-flattener, and nothing is sacrificed in the design of the plain Schmidt for the sake of improving performance when the field-flattener is added. The further question therefore arises: what improvement in performance results when the field-flattened Schmidt is designed as a single unit? This point is taken up in the second part of Section 5, where it is shown that, in systems of obstruction-ratio near to $\frac{1}{4}$, the performance of the field-flattened Schmidt is much improved by using a corrector plate with neutral zone close to the edge of the aperture.* Such a corrector plate would be undesirable in an $f/3$ plain Schmidt, even though the monochromatic aberrations of the systems are reduced by using it†, because of the chromatic difference of focus which it introduces. In the field-flattened Schmidt, this chromatic difference of focus can be used to balance the colour-error of the field-flattener and the result is a system whose aberrations are not appreciably larger than those of a plain Schmidt. They now include a small amount of coma, however, and this may prove a drawback in systems intended for positional measurements.

2. *Notation; the aberrations of the ordinary Schmidt camera; effect of plate-shift.*—Fig. 1 shows the optical system of the ordinary Schmidt camera. For the present we suppose that the aspheric surface of the corrector plate faces the spherical mirror and passes through its centre of curvature C. (x, y, z) are rectangular Cartesian coordinates with C as origin and with CA as z-axis. We

* F. E. Ross (1), p. 406, reached a similar conclusion by showing that the overall axial image-spreads are reduced when the neutral zone of the corrector plate is taken close to the edge of the aperture. His conclusion about the best compromise design for a convertible system must be accepted with reserve, however, since his argument takes no account of the properties of the off-axis images.

† See E. H. Linfoot and E. Wolf (3), pp. 755, 756.

take the aperture stop of the system to be the circle $x^2+y^2=H^2$ in the plane $z=0$, i.e. in the tangent plane to the aspheric surface. In an actual system of focal ratio $f/2{\cdot}5$ or longer, the aspheric surface lies everywhere within a few thousandths of an inch of this plane.

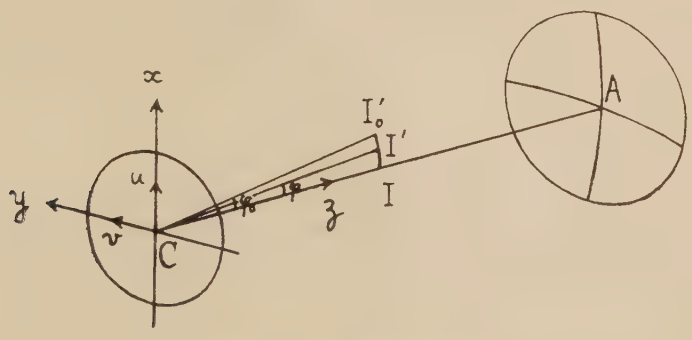

FIG. 1.—*The Schmidt camera.*

It is convenient to define the parameter

$$\mu = H/R, \tag{1}$$

where R is the radius of curvature of the mirror, and to introduce new variables u, v in the plane of the aperture stop by means of the equations

$$x = \mu R u, \qquad y = \mu R v; \tag{2}$$

then the boundary of the aperture stop is the circle $u^2+v^2=1$ and the equation of the aspheric surface can be written*

$$(n-1)z = \tfrac{1}{4}R\mu^4(r^4 - 2r_0^2r^2) + O(R\mu^6), \tag{3}$$

where $r^2=u^2+v^2$ and r_0 is the radius in (u, v)-coordinates of the neutral zone on the plate. In a practical case, the value of r_0 is usually about 0·8.

Rays entering the system parallel to the axis CA, and distant r_0H from it, meet the aspheric surface of the plate on the neutral zone and are therefore not deviated by the plate. After reflection by the mirror, they meet in a point I on CA, which we call the axial image-point of the camera. The sphere centred at C and passing through I is taken as the field-surface of the camera; its radius CI, which is also the focal length of the system, is given by the equation

$$\mathrm{CI} = \tfrac{1}{2}R(1-\mu^2r_0^2)^{-\frac{1}{2}}. \tag{4}$$

It follows that

$$\mu = \frac{H}{R} = \frac{H}{2\mathrm{CI}}(1-\mu^2r_0^2)^{-\frac{1}{2}} = \frac{1}{4F}(1-\mu^2r_0^2)^{-\frac{1}{2}}, \tag{5}$$

where F is the focal ratio of the system. (5) is often more useful in the approximate form

$$\mu = \frac{1}{4F}\{1+O(\mu^2)\}. \tag{6}$$

In an $f/3$ system $\mu=\frac{1}{12}$ to within an accuracy of the order of one per cent. Thus $\mu^2=1/144$ nearly, and we are justified in treating μ^2 as small in approximations.

* See Linfoot (**2**) for a derivation of the results quoted in this section. In some of the approximate formulae λ has here been replaced by μ, a legitimate procedure since, by equation (40) of (**2**), $\mu=\mu[1+O(\mu^2)]$.

If the system works over a field of angular diameter $2\phi_0$, the linear diameter of the field-surface is $2CI\sin\phi_0$ and the central obstruction-ratio

$$c=2F\sin\phi_0. \tag{7}$$

In Fig. 1, I_0' represents a point at the edge of the field-surface in the plane xCz; I' represents an intermediate field-point in the same plane, at an angular distance ϕ from the centre of the field. The parameter V, defined by the equation

$$V=\frac{\sin\phi}{\sin\phi_0}, \tag{8}$$

runs from zero at the centre of the field to 1 at its edge.

The point I' is, in third-order approximation, the image corresponding to an object point at infinity in the uz-plane, at an angular distance ϕ below the axis of the system. That is to say, the rays of a parallel pencil from this object point will eventually meet the field-surface in points whose distances from I' are $O(R\mu^5)$.

More precisely, the angular displacement-components from I', measured in seconds of arc, of a ray which enters through the point u, v of the aperture stop are respectively

$$\delta X', \delta Y' = K\mu^5\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right)\Phi^*(u, v, V;\ n') + O(K\mu^7), \tag{9}$$

where

$$K=648{,}000/\pi,$$

$$\Phi^*(u, v, V;\ n') = \left[\tfrac{1}{2}c^2V^2\left(\frac{1}{n}+u\frac{\partial}{\partial u}\right)+k_{nn'\mu}\right](r^4-2r_0^2r^2), \tag{10}$$

n' is the index of refraction of the plate for light of the wave-length under consideration, n that for the light in which the system is axially stigmatic and

$$k_{nn'\mu}=\frac{n'-n}{4\mu^2(n-1)}. \tag{11}$$

The value of $k_{nn'\mu}$ seldom exceeds 2 in a practical case. The "fifth-order aberration function"

$$R\mu^6\Phi^*(u, v, V;\ n') \tag{12}$$

measures, with error $O(R\mu^8)$, the distortions of the near-spherical wave-fronts converging to the point I'. u and v here play the part of parameters which identify the points of a wave-front by means of the intersections with the plane of the aperture stop of the rays through these points; and a positive value of Φ^* corresponds to a retardation of the wave-front.

From (3) the effect on the aberration-function of moving the plate a small distance $GRc^2\mu^2$ along the axis can be calculated. First suppose that the plate consists simply of the figuring layer, so that it can be treated as a plane lamina in our approximations. On the wave-fronts of a pencil of rays of n-light parallel to the axis, the retardation imposed by the plate is then*

$$\tfrac{1}{4}R\mu^4(r^2-r_0^2)^2+\tfrac{1}{8}R\mu^6P(r^2)+O(R\mu^8), \tag{13}$$

where $P(r^2)$ stands for the polynomial $(3r^2+2r_0^2)(r^2-r_0^2)^2$.

If the plate is moved a distance $GRc^2\mu^2$ along the axis towards the mirror, the aperture stop remaining in its original position, then a ray of an off-axis pencil which passes through the point (u, v) of the aperture stop now meets the

* This follows from equation (38) of (2).

plate in the point $(u+GRc^2\mu^2\sin\phi, V)=(u+2Gc^3\mu^2V, v)$. The retardation imposed by the plate at this point is obtained on replacing r^2 by r_V^2 in (13), where

$$r_V^2=(u+2Gc^3\mu^2V)^2+v^2$$
$$=r^2+4Gc^3\mu^2V+O(\mu^4). \tag{14}$$

Hence the change in the aberration-function is

$$\tfrac{1}{4}R\mu^4\{(r_V^2-r_0^2)^2-(r^2-r_0^2)^2\}+\tfrac{1}{8}R\mu^6\{P(r_V^2)-P(r^2)\}+O(R\mu^8)$$
$$=2RGc^3\mu^6Vu(r^2-r_0^2)+O(R\mu^8). \tag{15}$$

The term $-2RGc^3\mu^6Vur_0^2$ represents a simple tilt of the whole wave-front, causing a bodily displacement of the image through an angular distance $2Gc^3\mu^5Vr^2$ towards the centre of the field, i.e. a change in the equivalent focal length of the system by a factor $1-Gc^2\mu^4r_0^2$. The other main term, namely $2RGc^3\mu^6Vur^2$, measures (with error $O(R\mu^8)$) the primary coma which the plate-shift contributes to the aberration-function.

A change in the wave-length of the light, bringing with it a change $O(\mu^2)$ in the value of the refractive index, introduces a factor $1+O(\mu^2)$ into the expressions (13) and (15). Since the effect of this factor can be absorbed in the error term, it follows that the coma-contribution is independent of wave-length to the present order of approximation.

If the aperture stop is moved with the plate, the expressions for the aberration-function and for $\delta X, \delta Y$ remain unchanged in form. For if (u', v') are coordinates in the new aperture stop, measured parallel to (u, v), then evidently

$$u'=u+O(\mu^2), \qquad v'=v+O(\mu^2) \tag{16}$$

and on substituting for u, v in (9), (10) we obtain new expressions whose leading terms have the same algebraical form as before, but with u', v' in place of u, v. Since the boundary of the new aperture stop is the circle $u'^2+v'^2=1$, it follows that moving the aperture stop through a distance $O(R\mu^2)$ has left the fifth-order aberrations unaffected. The assumption that the plate-thickness is negligible can now be dropped, since the addition of a plane-parallel plate in front of the entry pupil does not affect the aberrations of the system.

When the plate is turned with its aspheric side away from the mirror, all the above results continue to hold, provided that in measuring the distance of the aspheric surface or the aperture stop from the pole A of the mirror, that part of the distance which lies in glass be taken with a factor $1/n$.

3. *Aberrations of the field-flattened Schmidt camera.*—Fig. 2 shows the system with the field-flattener inserted. The convex front face of the field-flattener

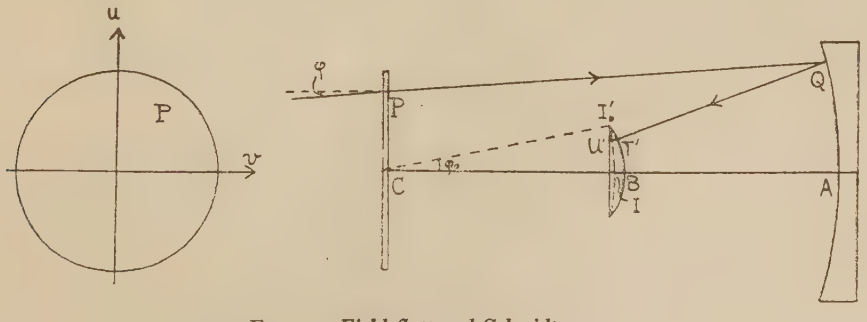

FIG. 2.—*Field-flattened Schmidt camera.*

has radius of curvature $\frac{n_1 - 1}{n_1}$ CI, where n_1 is the index of refraction of the glass for light of a selected wave-length; its sharp edge coincides with the edge $\phi = \phi_0$ of the original field-surface, and its flat back face forms the new field-surface. In practice, we should use a lens which is not sharp at the edge, placed with its back face not quite in contact with the photographic plate; the effect of these changes on the aberrations will be considered later. It will appear later that performance is improved by moving the corrector plate a short distance along the axis towards the mirror. For the present, we consider the system without this refinement.

PQ represents a ray, originating in the distant object point at an angular distance ϕ below the axis of the system, which passes through the point $P = (u, v)$ in the plane of the aperture stop. After reflection by the mirror at the point Q and refraction by the field-flattener at the point T′, this ray meets the focal plane in U′. To calculate the aberrations of the system in n'-light, we first suppose the rays from the concave mirror to converge exactly on to points of the virtual image surface II′ (see Fig. 3); that is to say, we disregard for the time being the aberrations

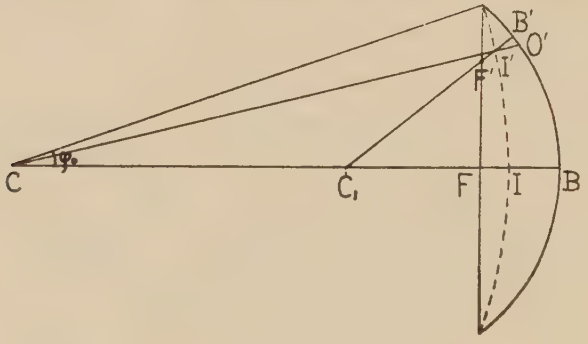

FIG. 3.

with which the converging pencils are loaded when they arrive at the field-flattener. Since the diameter of the field-flattener is $2\mathrm{CI}_0 \sin\phi_0 = O(R\mu)$, its thickness is $O(R\mu^2)$ and the lateral aberrations introduced when the virtual object point I′ is imaged in the convex surface of the lens are $O(R\mu^5)$. Moreover the Gauss image of I′ in the convex surface is at a distance $O(R\mu^4)$ from the adopted focal plane FF′, since the field-flattener has been given the Petzval curvature 1/CI. Hence the image-spreads on the adopted field-surface due to aberrations introduced by the field-flattener are $O(R\mu^5)$; that is, they are of the same order of magnitude as the image-spreads in the ordinary Schmidt camera. The same is true, as will be shown below, of the chromatic aberrations introduced by the field-flattener, provided that the refractive index of the field-flattener varies by only $O(\mu^2)$ throughout the effective spectral range.

It follows that, to obtain for the aberrations of the field-flattened system approximate formulae of the same order of accuracy as those quoted above for the ordinary Schmidt camera, it is sufficient to calculate separately the aberrations introduced by the field-flattener and by the rest of the system and to add the results, the cross-terms being of the order $R\mu^7$.

In Fig. 4, TU shows the path of a ray, originally proceeding towards the axial image point I, after refraction at the first surface of the field-flattener. U is the

intersection of this ray with the back surface and V is the point in which TU produced meets the axis. Let n_1 be the refractive index of the lens for the light in which Petzval curvature is compensated, so that

$$r_1 = \frac{n_1 - 1}{n_1} \mathrm{CI} = \frac{n_1 - 1}{n_1} \frac{R}{2} [1 + O(\mu^2)], \tag{17}$$

where CI is the distance of I from the centre of curvature of the mirror. Let n_1' be the refractive index of the lens in the light under consideration; we suppose, what is true in a practical case, that

$$n_1' - n_1 = O(\mu^2). \tag{18}$$

We write

$$\theta = \mu\Theta, \quad \theta' = \mu\Theta', \quad l = \mu^2 L, \quad y = \mu^3 Y, \quad \gamma = \mu^3 \Gamma; \tag{19}$$

then, by Fig. 4, the capital letters denote positive quantities which are $O(1)$ and

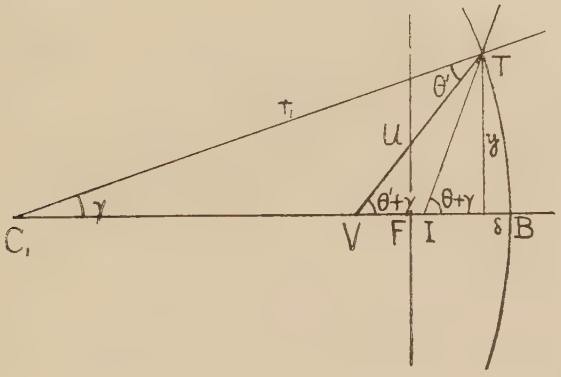

FIG. 4.

whose ratios are all $O(1)$. Thus the power of μ in each term of our expressions will give an indication of its order of magnitude. Evidently

$$\delta = O(y^2) = O(\mu^6).$$

From Fig. 4

$$\tan(\theta + \gamma) = \frac{y}{l - \delta},$$

$$\theta + \gamma + \tfrac{1}{3}(\theta + \gamma)^3 + O(\mu^5) = \frac{y}{l} + O\left(\frac{y\delta}{l^2}\right),$$

$$\mu\Theta + \mu^3\Gamma + \tfrac{1}{3}\mu^3\Theta^3 = \mu Y/L + O(\mu^5),$$

$$\Theta = \frac{Y}{L} - \mu^2\left(\Gamma + \frac{1}{3}\frac{Y^3}{L^3}\right) + O(\mu^4).$$

Now $\sin\gamma = y/r_1$; hence $\Gamma = \dfrac{Y}{r_1} + O(\mu^6)$ and we obtain

$$\Theta = \frac{Y}{L} - \left(\frac{Y}{r_1} + \frac{1}{3}\frac{Y^3}{L^3}\right)\mu^2 + O(\mu^4). \tag{20}$$

By Snell's law

$$n_1' \sin\theta' = \sin\theta,$$

$$\theta' = \frac{1}{n_1'}\left[\theta - \tfrac{1}{6}\theta^3\left(1 - \frac{1}{n_1'^2}\right)\right] + O(\theta^5),$$

$$\Theta' = \frac{1}{n_1'}\Theta\left[1 - \tfrac{1}{6}\mu^2\Theta^2\left(1 - \frac{1}{n_1'^2}\right) + O(\mu^4)\right]. \tag{21}$$

Hence

$$\begin{aligned}\mu\cot\theta' &= \frac{\mu}{\theta'}[1 - \tfrac{1}{3}\theta'^2 + O(\mu^4)] \\ &= \frac{1}{\Theta'}[1 - \tfrac{1}{3}\mu^2\Theta'^2 + O(\mu^4)] \\ &= n_1'\frac{L}{Y}\left[1 + \left(\frac{L}{r_1} + \frac{n_1'^2 - 1}{2n_1'^2}\frac{Y^2}{L^2}\right)\mu^2 + O(\mu^4)\right]. \end{aligned} \tag{22}$$

Again, from Fig. 4,

$$\frac{r_1}{C_1V} = \frac{\sin(\theta'+\gamma)}{\sin\theta'} = \cos\gamma + \cot\theta'\sin\gamma$$

$$= 1 + \frac{n_1'L\Gamma}{Y}\mu^2 + \frac{n'L\Gamma}{Y}\left(\frac{L}{r_1} + \frac{n_1'^2 - 1}{2n_1'^2}\frac{Y^2}{L^2}\right)\mu^4 + O(\mu^6);$$

$$C_1V = r_1\left\{1 - \frac{n_1'L\Gamma}{Y}\mu^2 + \left[\frac{n_1'^2L^2\Gamma^2}{Y^2} - \frac{n_1'L\Gamma}{Y}\left(\frac{L}{r_1} + \frac{n_1'^2 - 1}{2n_1'^2}\frac{Y^2}{L^2}\right)\right]\mu^4 + O(\mu^6)\right\}$$

$$= r_1\left[1 - \frac{n_1'L}{r_1}\mu^2 + \frac{n_1'(n_1'-1)L^2}{r_1^2}\mu^4 - \frac{n_1'^2 - 1}{2n_1'}\frac{Y^2}{Lr_1}\mu^4 + O(\mu^6)\right] \tag{23}$$

$$= r_1\left[1 - \frac{n_1'l}{r_1} + \frac{n_1'(n_1'-1)}{r_1^2}l^2 - \frac{n_1'^2 - 1}{2n_1'}\frac{y^2}{lr_1} + O(\mu^6)\right],$$

$$BV = r_1 - C_1V = n_1'l - \frac{n_1'(n_1'-1)}{r_1}l^2 + \frac{n_1'^2 - 1}{2n_1'}\frac{y^2}{l} + O(r_1\mu^6) \tag{24}$$

The first part of this expression, independent of y, is the expansion as far as terms of order μ^4 of the distance BF_0 of the Gaussian image F_0 from B. The remainder therefore gives an approximate expression for the longitudinal spherical aberration.

In the case of an off-axis pencil, shown in Fig. 5, we make use of an auxiliary optic axis, $C_1I'B'$, which no longer coincides with the principal ray $CI'O'$ of the incident pencil. An arbitrary ray of the pencil converging to the virtual object point I′ meets the convex surface in T′ and is refracted into the ray T′U′V′, which meets the back surface of the lens in U′ and the auxiliary axis C_1B' in V′. (Fig. 5 illustrates the special case of a meridional ray.) y', l' are respectively the distance of T′ from the auxiliary axis $C_1I'B'$ and the distance B′I′ of the virtual object point inside B′.

We first need expressions for $l' = B'I'$ and for $B'F'$. From the figure,

$$B'I' = r_1 - C_1I'; \qquad CI' = CI,$$

$$\begin{aligned} C_1I'^2 &= CC_1^2 + CI'^2 - 2CC_1 . CI'\cos\phi \\ &= 2CI(CI - r_1 + l)(1 - \cos\phi) + (r_1 - l)^2. \end{aligned} \tag{25}$$

Now

$$l = BI = FB - FI = r_1(1 - \cos\psi_0) - CI(1 - \cos\phi_0), \tag{26}$$

where ϕ_0 and ψ_0 are the values of ϕ, ψ at the edge of the field;

$$\sin\psi_0 = \frac{CI'_0}{C_1I'_0}\sin\phi_0 = \frac{n_1}{n_1-1}\sin\phi_0, \tag{27}$$

$$\sin\phi_0 = 2c\lambda = 2c\mu(1-\tfrac{1}{2}r_0^2\mu^2)+O(\mu^5). \tag{28}$$

On making use of (4), (17), (27) and (28), equation (26) gives

$$l = r_1\frac{2c^2\mu^2n_1}{(n_1-1)^2}\left[1-r_0^2\mu^2+\frac{3n_1^2-3n_1+1}{(n_1-1)^2}c^2\mu^2+O(\mu^4)\right],$$

$$(r_1-l)^2 = r_1^2\left\{1-\frac{4n_1c^2\mu^2}{(n_1-1)^2}\left[1-\left(r_0^2-\frac{3n_1-1}{n_1-1}c^2\right)\mu^2\right]+O(\mu^6)\right\}. \tag{29}$$

Also

$$1-\cos\phi = \tfrac{1}{2}\sin^2\phi[1+\tfrac{1}{4}\sin^2\phi]+O(\mu^6)$$
$$= 2c^2\mu^2V^2(1-r_0^2\mu^2+c^2\mu^2V^2)+O(\mu^6), \tag{30}$$

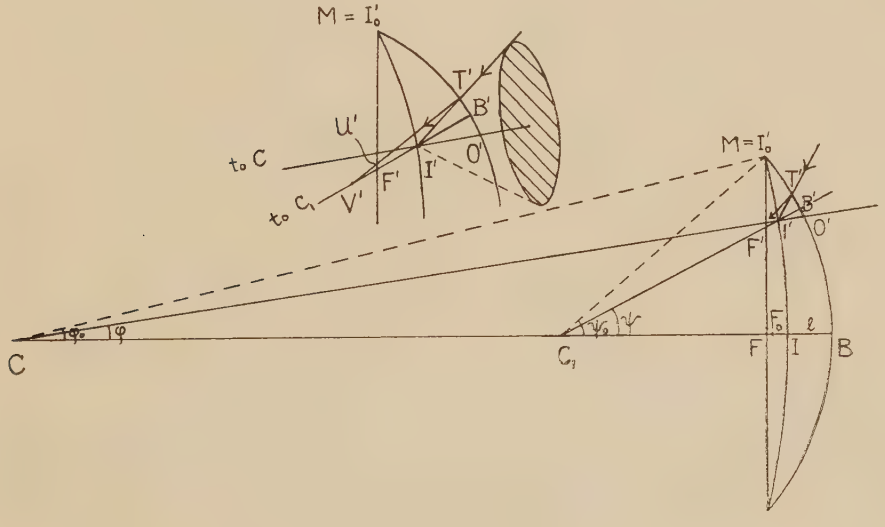

FIG. 5.

where V is defined by (8). Hence (25) gives

$$C_1I'^2 = r_1^2\left\{1-\frac{4n_1}{(n_1-1)^2}c^2\mu^2(1-V^2)\left[1-r_0^2\mu^2+c^2\mu^2\left(\frac{3n_1-1}{n_1-1}+V^2\right)\right]+O(\mu^6)\right\},$$

$$C_1I' = r_1\left\{1-\frac{2n_1}{(n_1-1)^2}c^2\mu^2(1-V^2)\left[1-r_0^2\mu^2+\right.\right.$$
$$\left.\left.+c^2\mu^2\left(\frac{3n_1^2-3n_1+1}{(n_1-1)^2}+\frac{n_1^2-3n_1+1}{(n_1-1)^2}V^2\right)\right]+O(\mu^6)\right\};$$

and we obtain finally

$$l' = B'I' = r_1-C_1I' = r_1\left\{\frac{2n_1c^2\mu^2}{(n_1-1)^2}(1-V^2)\left[1-r_0^2\mu^2+c^2\mu^2\left(\frac{3n_1^2-3n_1+1}{(n_1-1)^2}+\right.\right.\right.$$
$$\left.\left.\left.+\frac{n_1^2-3n_1+1}{(n_1-1)^2}V^2\right)\right]+O(\mu^6)\right\}. \tag{31}$$

Next,

$$B'F' = r_1 - C_1F', \qquad C_1F' = C_1F \sec\psi. \tag{32}$$

Here

$$C_1F = r_1 - BF = r_1\left[1 - \frac{2c^2\mu^2 n_1^2}{(n_1-1)^2}\left(1 - r_0^2\mu^2 + \frac{n_1^2c^2\mu^2}{(n_1-1)^2}\right) + O(\mu^2)\right], \tag{33}$$

by (27) and (28), while

$$\sin^2\psi = \frac{CI'^2}{C_1I^2}\sin^2\phi = \frac{4n_1^2c^2\mu^2V^2}{(n_1-1)^2}\left[1 - r_0^2\mu^2 + \frac{4n_1c^2\mu^2}{(n_1-1)^2}(1-V^2) + O(\mu^4)\right],$$

$$\sec\psi = 1 + \tfrac{1}{2}\sin^2\psi + \tfrac{3}{8}\sin^4\psi + O(\mu^6)$$

$$= 1 + \frac{2n_1^2c^2\mu^2V^2}{(n_1-1)^2}\left[1 - r_0^2\mu^2 + \frac{4n_1c^2\mu^2}{(n_1-1)^2}(1-V^2)\right] + 6\frac{n_1^4c^4\mu^4V^4}{(n_1-1)^4} + O(\mu^6). \tag{34}$$

From (32), (33), (34)

$$C_1F' = r_1\left[1 - \frac{2n_1^2}{(n_1-1)^2}c^2\mu^2(1-V^2)\right.$$
$$\left.\times\left\{1 - r_0^2\mu^2 + \frac{c^2\mu^2n_1^2}{(n_1-1)^2}(1+3V^2) - \frac{4c^2\mu^2n_1}{(n_1-1)^2}V^2\right\} + O(\mu^6)\right],$$

$$B'F' = r_1\frac{2n_1^2}{(n_1-1)^2}c^2\mu^2(1-V^2)$$
$$\times\left\{1 - r_0^2\mu^2 + \frac{c^2\mu^2n_1^2}{(n_1-1)^2}(1+3V^2) - \frac{4n_1}{(n_1-1)^2}c^2\mu^2V^2\right\} + O(r_1\mu^6), \tag{35}$$

(31) and (35) are the required expressions for l' and for $B'F'$.

The next stage is to calculate $B'V'$. Let ϖ be a plane drawn through O' perpendicular to CO' and let x'', y'' be coordinates in this plane, $O'x''$ being in the plane BCO' and $O'y''$ at right angles to it. Let T'', B'' be the perpendicular projections of T', B' on the plane ϖ. (See Fig. 6.)

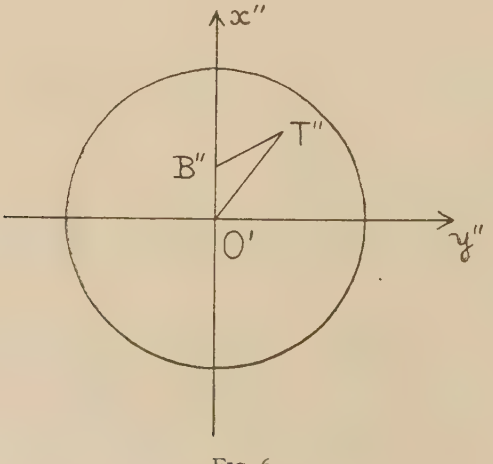

Fig. 6.

For a ray T′I′ (not necessarily a meridional ray) which entered the system through the point (u, v) of the aperture stop, T″ has coordinates

$$(x'', y'') = (u, v)\frac{H}{\frac{1}{2}R}\mathrm{O'I'}[1 + O(\mu^2)]$$

$$= (u, v)2\mu\mathrm{B'I'}[1 + O(\mu^2)]$$

$$= (u, v)r_1\frac{4n_1c^2\mu^3}{(n_1-1)^2}(1 - V^2)[1 + O(\mu^2)], \tag{36}$$

by (31). Now

$$\mathrm{B''O'} = \mathrm{B'O'}[1 + O(\mu^2)] = \mathrm{B'I'}(\psi - \phi)[1 + O(\mu^2)]$$

$$= r_1\frac{2n_1c^2\mu^2}{(n_1-1)^2}(1 - V^2)(\sin\psi - \sin\phi)[1 + O(\mu^2)]$$

$$= r_1\frac{2n_1c^2\mu^2}{(n_1-1)^2}(1 - V^2)\left(\frac{\mathrm{CI'}}{\mathrm{C_1I'}} - 1\right)2c\lambda V[1 + O(\mu^2)]$$

$$= r_1\frac{4n_1c^3\mu^3}{(n_1-1)^3}V(1 - V^2)[1 + O(\mu^2)] \tag{37}$$

and

$$y' = \mathrm{I'B'}\sin\mathrm{T'I'B'}\,.\,[1 + O(\mu^2)]$$

$$= \mathrm{T''B''}[1 + O(\mu^2)]$$

$$= r''[1 + O(\mu^2)], \tag{38}$$

where

$$r''^2 = \mathrm{T''B''^2} = \left(x'' - r_1\frac{4n_1c^3\mu^3}{(n_1-1)^3}V(1 - V^2)\right)^2 + y''^2$$

$$= r_1^2\left(\frac{4n_1c^2\mu^3}{(n_1-1)^2}\right)^2(1 - V^2)^2\left[\left(u - \frac{cV}{n_1-1}\right)^2 + v^2\right].[1 + O(\mu^2)], \tag{39}$$

and, by (31),

$$\frac{r''^2}{l'^2} = 4\mu^2\left[\left(u - \frac{cV}{n_1-1}\right)^2 + v^2\right].[1 + O(\mu^2)]. \tag{40}$$

Hence by (24), applied with the auxiliary axis $\mathrm{C_1I'}$ in place of $\mathrm{C_1I}$ and with l', y' in place of l, y,

$$\mathrm{B'V'} = n_1'l' - \frac{n_1'(n_1'-1)}{r_1}l'^2 + \frac{n_1'^2-1}{2n_1'}\frac{y'^2}{l'} + O(r_1\mu^6)$$

$$= n_1l' - \frac{n_1(n_1-1)}{r_1}l'^2 + \frac{n_1^2-1}{2n_1}\frac{y'^2}{l'} + (n_1'-n_1)l + O(r_1\mu^6), \tag{41}$$

in virtue of the hypothesis (18),

$$= r_1\frac{2n_1^2c^2\mu^2}{(n_1-1)^2}(1 - V^2)\left\{1 - r_0^2\mu^2 + c^2\mu^2\left(\frac{3n_1^2-3n_1+1}{(n_1-1)^2} + \frac{n_1^2-3n_1+1}{(n_1-1)^2}V^2\right)\right.$$

$$\left. - \frac{2n_1}{n_1-1}c^2\mu^2(1 - V^2) + \frac{n_1^2-1}{2n_1^2}4\mu^2\left[\left(u - \frac{cV}{n_1-1}\right)^2 + v^2\right] + \frac{n_1'-n}{n_1}\right\} + O(r_1\mu^6), \tag{42}$$

by (38), (39), (40). This is the desired expression for B′V′.

From (35) and (42),

$$\mathrm{F'V'}=\mathrm{B'V'}-\mathrm{B'F'}=r_1\frac{2n_1^2c^2\mu^4}{(n_1-1)^2}(1-V^2)\left\{2\frac{n_1^2-1}{n_1^2}\left[\left(u-\frac{cV}{n_1-1}\right)^2+v^2\right]\right.$$
$$\left.-\frac{c^2}{n_1-1}(1+V^2)+\frac{4(n_1-1)}{n_1}k_{n_1n_1'\mu}\right\}+O(r_1\mu^6), \tag{43}$$

where

$$k_{n_1n_1'\mu}=\frac{n_1'-n_1}{4\mu^2(n_1-1)}=O(1) \tag{44}$$

in a practical case.

Now introduce deviation-coordinates (ξ',η') in the adopted focal plane, i.e. the back face of the lens, taking as origin the intersection $\mathrm{F'}\left(\frac{cV}{n_1-1},0\right)$ of the ray B'I' (for which $u=\frac{cV}{n_1-1}$, $v=0$) with this plane. Since the semi-angle in glass of the whole cone of rays is

$$\frac{H}{n_1f}[1+O(\mu^2)]=\frac{2\mu}{n_1}[1+O(\mu^2)],$$

we obtain from (43)

$$\xi'+i\eta'=\frac{2\mu}{n_1}\left(u-\frac{cV}{n_1-1}+iv\right)\mathrm{F'V'}[1+O(\mu^2)]$$
$$=\frac{4n_1}{(n_1-1)^2}r_1c^2\mu^5(1-V^2)\left(u-\frac{cV}{n_1-1}+iv\right)$$
$$\times\left\{\frac{2(n_1^2-1)}{n_1^2}\left[\left(u-\frac{cV}{n_1-1}\right)^2+v^2\right]-\frac{c^2}{n_1-1}(1+V^2)+\frac{4(n_1-1)}{n_1}k_{n_1n_1'\mu}\right\}$$
$$=2R\mu^5c^2(1-V^2)\left(re^{iE}-\frac{cV}{n_1-1}\right)$$
$$\times\left\{\frac{2(n_1+1)}{n_1^2}\left(r^2-2r\cos E\frac{cV}{n_1-1}+\frac{c^2V^2}{(n_1-1)^2}\right)-\frac{c^2}{(n_1-1)^2}(1+V^2)\right.$$
$$\left.+\frac{4}{n_1}k_{n_1n_1'\mu}\right\}+O(R\mu^7),$$

where $u+iv=re^{iE}$,

$$=2R\mu^5c^2(1-V^2)\left\{\frac{2(n_1+1)}{n_1^2}\left[r^3e^{iE}-r^2\frac{cV}{n_1-1}(2+e^{2iE})\right.\right.$$
$$\left.+r\frac{c^2V^2}{(n_1-1)^2}(2e^{iE}+e^{-iE})-\frac{c^3V^3}{(n_1-1)^3}\right]+\left(re^{iE}-\frac{cV}{n_1-1}\right)$$
$$\left.\times\left(\frac{4}{n_1}k_{n_1n_1'\mu}-\frac{c^2}{(n_1-1)^2}(1+V^2)\right)\right\}+O(R\mu^7). \tag{45}$$

The intersection $\mathrm{F'}(0,0)$ of the principal ray O'I' (for which $u=0,v=0$) with the focal plane is obtained on setting $r=0$ in (45). It follows that if we move the origin of the deviation-coordinates to $\mathrm{F'}(0,0)$ and denote the new coordinates by (ξ'',η''), the deviations are given by the equation

$$\xi''+i\eta''=2R\mu^5c^2(1-V^2)\left\{\frac{2(n_1+1)}{n_1^2}\left[r^3e^{iE}-r^2\frac{cV}{n_1-1}(2+e^{2iE})\right.\right.$$
$$\left.\left.+r\frac{c^2V^2}{(n_1-1)^2}(2e^{iE}+e^{-iE})\right]-re^{iE}\frac{c^2(1+V^2)}{(n_1-1)^2}+\frac{4}{n_1}re^{iE}k_{n_1n_1'\mu}\right\}+O(R\mu^7) \tag{46}$$

The corresponding angular deviation-components $\delta X''$ and $\delta Y''$, measured in seconds of arc, are obtained on dividing by $\pi/648{,}000$ times the focal length of the system; they are therefore given by the equation

$$\delta X''+i\delta Y''=4K\mu^5c^2(1-V^2)\left\{\frac{2(n_1+1)}{n_1^2}\left[r^3e^{iE}-r^2\frac{cV}{n_1-1}(2+e^{2iE})\right.\right.$$
$$\left.\left.+r\frac{c^2V^2}{(n_1-1)^2}(2e^{iE}+e^{-iE})\right]-\frac{c^2(1+V^2)}{(n_1-1)^2}re^{iE}+\frac{4}{n_1}k_{n_1n_1'\mu}re^{iE}\right\}+O(K\mu^7), \tag{47}$$

where $K=\dfrac{648{,}000}{\pi}$, as before. (47) can be written in the form

$$(\delta X'',\delta Y'')=K\mu^5\left(\frac{\partial\Phi'}{\partial u},\ \frac{\partial\Phi'}{\partial v}\right)+O(K\mu^7), \tag{48}$$

where

$$\Phi'=\Phi'(u,v,V;\ n_1')$$
$$=\frac{4c^2(1-V^2)}{n_1^2}\left\{\frac{n_1+1}{2}\left(1-\frac{cV}{n_1-1}\frac{\partial}{\partial u}\right)r^4+\left[2n_1k_{n_1n_1'\mu}-\frac{n_1^2}{2(n_1-1)^2}c^2(1+V^2)\right.\right.$$
$$\left.\left.+\frac{n_1+1}{(n_1-1)^2}c^2V^2\left(1+u\frac{\partial}{\partial u}\right)\right]r^2\right\}. \tag{49}$$

From (9), (10), (48), (49), the angular deviation-components $(\delta X,\delta Y)$ of the complete system, measured in seconds of arc, are given by the equation

$$(\delta X,\delta Y)=(\delta X',\delta Y')+(\delta X'',\delta Y'')+O(K\mu^7)$$
$$=K\mu^5\left(\frac{\partial}{\partial u},\frac{\partial}{\partial v}\right)\Psi(u,v,V)+O(K\mu^7), \tag{50}$$

where

$$\Psi(u,v,V)=\Phi^*(u,v,V;\ n')+\Phi'(u,v,V;\ n_1')$$
$$=\left[\tfrac{1}{2}c^2V^2\left(\frac{1}{n}+u\frac{\partial}{\partial u}\right)+k_{nn'\mu}\right](r^4-2r_0^2r^2)$$
$$+\frac{4c^2(1-V^2)}{n_1^2}\left\{\frac{n_1+1}{2}\left(1-\frac{cV}{n_1-1}\frac{\partial}{\partial u}\right)r^4+\left[2n_1k_{n_1n_1'\mu}\right.\right.$$
$$\left.\left.-\frac{n_1^2}{2(n_1-1)^2}c^2(1+V^2)+\frac{n_1+1}{(n_1-1)^2}c^2V^2\left(1+u\frac{\partial}{\partial u}\right)\right]r^2\right\}. \tag{51}$$

Here n, n_1 denote the refractive indices of corrector plate and field-flattener at that point of the spectrum for which the corrector plate exactly compensates the primary spherical aberration of the mirror; n', n_1' denote their refractive indices in the light under consideration, and it is assumed that

$$n'=n+O(\mu^2),\qquad n_1'=n_1+O(\mu^2).$$

The function $R\mu^6\Psi(u,v,V)$, which may be called the *fifth-order aberration function* of the field-flattened Schmidt camera, describes the distortion of the near-spherical wave-fronts converging towards the image point. Its value measures, with error $O(R\mu^8)$, the optical distance through which the point of parameter (u,v) on the wave-front is pushed backward from the corresponding spherical wave-front centred at the image point F′(0, 0). Such approximate

aberration-functions not only allow the ray-theoretic properties of an optical system to be expressed in a compact form, but can also form a basis for the calculation of its diffraction properties by means of Huygens' principle.

For example, in an $f/2{\cdot}5$ Schmidt camera of 20 inches aperture, $\mu=\frac{1}{10}$ and $R=100$ inches, so that $R\mu^8=10^{-6}$ inch $=\frac{1}{20}$ wave-length, approximately. Thus the fifth-order aberration function can safely be used for the computation of the diffraction images of the system.

In a 24-inch $f/3$ system working over a 5° field, the diameter of the field-flattener is six inches and its thickness decreases from about $\frac{3}{16}$ inch at the centre to zero at the edge. Such a lens would be too fragile for practical use, quite apart from the difficulty of grinding and polishing it. It would be better to use a lens whose thickness decreases from about $\frac{5}{16}$ inch at the centre to about $\frac{1}{8}$ inch at the edge.

The above analysis can be applied to this more practical type of field-flattener by treating it as the central portion of a sharp-edged lens originally designed to flatten a field of diameter $\sqrt{\frac{5}{3}}$ times that actually in question. We need only replace c, V in (51) by $c\sqrt{\frac{5}{3}}$, $V\sqrt{\frac{3}{5}}$ respectively to obtain the fifth-order aberration function of the new system, namely $R\mu^6$ times the function

$$\Psi(u,v,V)=\left[\tfrac{1}{2}c^2V^2\left(\frac{1}{n}+u\frac{\partial}{\partial u}\right)+k_{nn'\mu}\right](r^4-2r_0^2r^2)$$
$$+\frac{4c^2}{n_1^2}(\tfrac{5}{3}-V^2)\left\{\frac{n_1+1}{2}\left(1-\frac{cV}{n_1-1}\frac{\partial}{\partial u}\right)r^4+\left[2n_1k_{n_1n_1'\mu}\right.\right.$$
$$\left.\left.-\frac{n_1^2}{2(n_1-1)^2}c^2(\tfrac{5}{3}+V^2)+\frac{n_1+1}{(n_1-1)^2}c^2V^2\left(1+u\frac{\partial}{\partial u}\right)\right]r^2\right\}. \qquad (52)$$

It is inconvenient in practice to have the photographic plate actually in contact with the back surface of the field-flattener, since this eventually results in the latter becoming scratched. The inconvenience can be avoided by moving forward the back surface of the field-flattener towards its convex surface, so as to reduce its thickness by an amount $O(R\mu^4)$. This change introduces aberration-displacements of only $O(R\mu^7)$ in the new focal plane, which is now in air and slightly in front of the old one. Its effects on the aberrations are therefore negligible in the present approximation.

In a 24-inch $f/3$ system, $R\mu^4=\frac{1}{144}$ inch and if we cut away the back face of the field-flattener to a depth of $5R\mu^4=\frac{1}{30}$ inch, approximately, the aberrations are not changed by more than a few per cent. There is then a clearance of about $\frac{1}{50}$ inch between the emulsion layer and the back surface of the lens.

4. *Balancing the aberrations.*—From equation (52) we see that the only asymmetrical contribution to the fifth-order aberrations of the field-flattened Schmidt camera is that represented by the term

$$-2R\mu^6c^3V(\tfrac{5}{3}-V^2)\frac{n_1+1}{n_1^2(n_1-1)}\frac{\partial}{\partial u}r^4$$
$$=-8R\mu^6c^3V(\tfrac{5}{3}-V^2)\frac{n_1+1}{n_1^2(n_1-1)}ur^2. \qquad (53)$$

This coma, being proportional to $V(\frac{5}{3}-V^2)$, vanishes at the centre of the field and is greatest at $V=\frac{1}{3}\sqrt{5}$. Its amount is very small in practice; in a 24-inch $f/3$ system working over a 5° field ($c=\frac{1}{4}$), the coma-spread is nowhere as large as 0·57 second of arc.

A still smaller coma-value can be obtained by moving the corrector-plate a distance

$$\frac{11}{3}\frac{n_1+1}{n_1^2(n_1-1)}Rc^2\mu^2 \quad (=8Rc^2\mu^2, \text{ nearly}) \tag{54}$$

along the axis towards the mirror. For, by (15), this adds a term

$$\tfrac{22}{3}R\frac{n_1+1}{n_1^2(n_1-1)}c^3\mu^6 Vur^2 \tag{55}$$

to the aberration-function, in which the coma-term (53) therefore becomes

$$-8R\mu^6c^3V(\tfrac{3}{4}-V^2)\frac{n_1+1}{n_1^2(n_1-1)}ur^2. \tag{56}$$

Since the maximum value of $|V(\frac{3}{4}-V^2)|$ in the range $0\leqq V\leqq 1$ is 0·25, while that of $V(\frac{5}{3}-V^2)$ is $\frac{10}{27}\sqrt{5}=0{\cdot}83$, the plate-shift has reduced the maximum coma-spread to less than one-third of its previous value.

In a 24-inch $f/3$ system, working over a 5° field, the plate must be shifted 0·5 inch towards the mirror to balance the coma, which then nowhere exceeds 0·2 second of arc.

In the coma-balanced systems, (52) is replaced by the equation

$$\begin{aligned}\Psi(u,v,V)=&\left[\tfrac{1}{2}c^2V^2\left(\frac{1}{n}+u\frac{\partial}{\partial u}\right)+k_{nn'\mu}\right](r^4-2r_0^2r^2)\\&+\frac{4c^2}{n_1^2}(\tfrac{5}{3}-V^2)\left\{\frac{n_1+1}{2}r^4+\left[4n_1k_{n_1n'_1\mu}-\frac{n_1^2c^2}{(n_1-1)^2}(\tfrac{5}{3}+V^2)\right.\right.\\&\left.\left.+\frac{n_1+1}{(n_1-1)^2}c^2V^2\left(1+u\frac{\partial}{\partial u}\right)\right]r^2\right\}-8c^3V(\tfrac{3}{4}-V^2)\frac{n_1+1}{n_1^2(n_1-1)}ur^2.\end{aligned} \tag{57}$$

We have next to consider whether any appreciable improvement in performance can be obtained by small adjustments of the field-curvature of the system and of the focal setting. (The field-curvature adjustment could be most conveniently made by means of a small change in r_1, the radius of curvature of the convex surface of the field-flattener.)

In the focal plane so far adopted, the fifth-order angular deviation-components are, in seconds of arc,

$$(\delta X^*,\delta Y^*)=K\mu^5\left(\frac{\partial\Psi}{\partial u},\frac{\partial\Psi}{\partial v}\right),$$

where $\Psi=\Psi(u,v,V)$ is defined by (57). A moving forward of the focal plane by $\frac{1}{2}\epsilon f\mu^4$ (in air) adds $K\epsilon\mu^5(u,v)$ to these angular deviation-components.

We define the "effective image radius" at a particular point of the field as the square root of the mean value

$$\frac{1}{\pi}\iint_{u^2+v^2\leqq 1}[(\delta X^*)^2+(\delta Y^*)^2]\,du\,dv,$$

and the "best focus" as the focal setting which minimizes the effective image radius.* Thus to obtain the best focus at any point of the field we have to choose

* See (3), p. 754, for a discussion of the reasons for adopting this definition.

ϵ so as to minimize the expression

$$\iint_{u^2+v^2\leqq 1}\left\{\left(\frac{\partial\Psi}{\partial u}+\epsilon u\right)^2+\left(\frac{\partial\Psi}{\partial v}+\epsilon v\right)^2\right\}du\,dv \tag{58}$$

at this point. Let $\epsilon_0=\epsilon_0(V)$ denote this minimizing value of ϵ; then ϵ_0 satisfies the equation

$$-\epsilon_0\iint_{u^2+v^2\leqq 1}(u^2+v^2)\,du\,dv=\iint_{u^2+v^2\leqq 1}\left(u\frac{\partial\Psi}{\partial u}+v\frac{\partial\Psi}{\partial v}\right)du\,dv$$

or

$$-\frac{\pi}{2}\epsilon_0=J(\Psi), \tag{59}$$

where $J(f)$ stands for the expression

$$\iint_{u^2+v^2\leqq 1}\left(u\frac{\partial}{\partial u}+v\frac{\partial}{\partial v}\right)f\,du\,dv. \tag{60}$$

We easily find

$$\left.\begin{aligned}&J(r^4)=\tfrac{4}{3}\pi,\quad J\left(u\frac{\partial}{\partial v}r^2\right)=\pi,\\&J(r^4-2r_0^2r^2)=(\tfrac{4}{3}-2r_0^2)\pi,\\&J\left[u\frac{\partial}{\partial u}(r^4-2r_0^2r^2)\right]=(\tfrac{8}{3}-2r_0^2)\pi,\end{aligned}\right\} \tag{61}$$

and it follows from (57), (59), (60), (61) that

$$\begin{aligned}\epsilon_0(V)=2(2r_0^2-\tfrac{4}{3})&\left(k_{nn'\mu}+\frac{2n_1+1}{2n_1}c^2V^2\right)-2r_0^2c^2V^2\\&-\frac{8c^2}{n_1^2}(\tfrac{5}{3}-V^2)\left[\tfrac{2}{3}(n_1+1)+4n_1k_{n_1n'_1\mu}-\tfrac{5}{3}\frac{n_1^2c^2}{(n_1-1)^2}\right]\\&-\frac{8c^4}{n_1^2}V^2(\tfrac{5}{3}-V^2)\left(\frac{3}{(n_1-1)^2}-1\right).\end{aligned} \tag{62}$$

The value, in seconds of arc, of the effective image radius at this focal setting is $1+O(\mu^2)$ times the square root of

$$K\mu^5\left[\iint\left\{\left(\frac{\partial\Psi}{\partial u}\right)^2+\left(\frac{\partial\Psi}{\partial v}\right)^2\right\}du\,dv-\frac{\pi}{2}\epsilon_0^2\right].$$

Equation (62) specifies the best-field surface in each separate wave-length, measured from the originally chosen field-surface as reference plane. On the right of (62), the terms independent of V represent, to the present order of approximation, a simple shift of the focal plane; those in V^2 represent a residual curvature of the best field and those in V^4 a deviation of the best-field surface from the spherical form. We next examine these residuals to see whether any appreciable improvement in the performance of the system can be obtained by adjusting the value of r_1.

We suppose that $a=2r_0^2$ has the value which optimizes the axial colour-error of the original Schmidt camera in the sense of minimizing the effective radii of the axial colour-confusion patches, namely $a=\frac{4}{3}$*, and that $c=\frac{1}{4}$, $n_1=\frac{3}{2}$. In most astronomical Schmidt cameras, a, c and n_1 are near to these values.

* See (3), p. 756, footnote. The corresponding value of r_0 is $\sqrt{\frac{2}{3}}=0{\cdot}816$.

Then (62) gives

$$\epsilon_0(V) = -\tfrac{1}{12}V^2 - \tfrac{2}{9}(\tfrac{5}{3} - V^2)(\tfrac{35}{48} + 6k_{n_1 n_1' \mu}) - \tfrac{11}{72}V^2(\tfrac{5}{3} - V^2) \qquad (63)$$

and in particular, for $k_{n_1 n_1' \mu} = 0$, i. e. for $n_1' = n_1$,

$$\epsilon_0(V) = -0{\cdot}270 - 0{\cdot}176V^2 + 0{\cdot}153V^4. \qquad (64)$$

Fig. 7 shows the values of $\epsilon_0(V)$, in the case (63), for C, F and h-light, on the assumption that the field-flattener is of hard crown glass (Chance HC 519604) with

$$(n_1)_C = 1{\cdot}51637,\ n_1 = (n_1)_F = 1{\cdot}52496,\ (n_1)_h = 1{\cdot}53349.$$

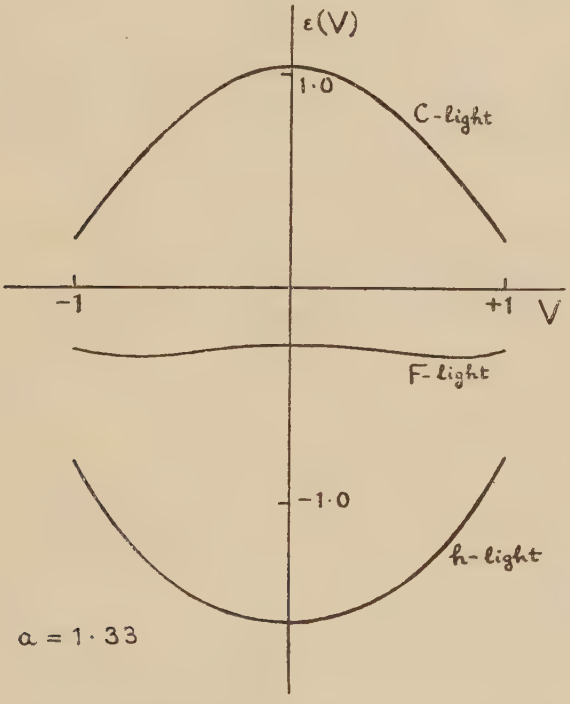

FIG. 7.—*Surfaces of best focus in the system with field-flattener added.*

It appears from Fig. 7 that a constant focus shift $\epsilon = -0{\cdot}30$ will improve the images in F-light and that there is no appreciable gain in performance to be obtained by changing the value of r_1. The choice $\epsilon = -0{\cdot}30$ corresponds to a focus shift of $-\frac{1}{2} \times 0{\cdot}30 f\mu^4$ in air, i. e. to an increase of $0{\cdot}15 f\mu^4$ in the air-gap between the photographic plate and the back surface of the field-flattener.* In a 24-inch $f/3$ system, this amounts to 0·0005 inch, which is within the practical focusing tolerance of an $f/3$ photographic telescope. However, because of the high theoretical performance of the system, it is necessary to take it into account in calculating the geometrical aberrations.

With the new choice of focal plane, the aberration-function becomes

$$R\mu^6[\Psi(u, v, V) + \tfrac{1}{2}\epsilon(u^2 + v^2)]$$
$$= R\mu^6[\Psi(u, v, V) - 0{\cdot}15(u^2 + v^2)], \qquad (65)$$

* The difference between $\frac{1}{2}\epsilon f\mu^4$ and $\frac{1}{2}\epsilon f\mu^4 \sec\phi$, being $O(f\mu^6)$, can be neglected in the present approximation.

where $\Psi(u, v, V)$ is given by (57), and the angular aberration-components, measured in seconds of arc, are

$$\delta X, \delta Y = K\mu^5 \left(\frac{\partial \Psi}{\partial u} - 0{\cdot}30u, \frac{\partial \Psi}{\partial v} - 0{\cdot}30v \right) + O(K\mu^7). \tag{66}$$

5. *An application.*—The general formulae (57) and (66) provide a convenient means of comparing the theoretical performances of the field-flattened Schmidt and the plain Schmidt at focal ratios such as are used in large astronomical cameras. As a typical example, we consider a 24-inch $f/3$ system working over a 5° field.

We suppose that the corrector plate has its neutral zone-radius $r_0 = 0{\cdot}816$ of the full aperture, and that when used without the field-flattener the system is axially stigmatic in F-light. The refractive indices in C, F and h-light of the corrector plate, supposed made of white plate glass, are taken to be

$$n_C = 1{\cdot}5189, \quad n_F = 1{\cdot}5278, \quad n_h = 1{\cdot}5366. \tag{67}$$

Since the system is assumed axially stigmatic in F-light, we have to set $n = n_F$ in the general formulae; we also have $\mu = \frac{1}{12}$, $c = \frac{1}{4}$, $a = \frac{4}{3}$.

We suppose the field-flattener to be of crown glass (Chance HC 519604) with

$$(n_1)_C = 1{\cdot}51637, \quad n_1 = (n_1)_F = 1{\cdot}52496, \quad (n_1)_h = 1{\cdot}53349. \tag{68}$$

Small variations in these constants would not appreciably alter the properties of the system.

Finally, we suppose the field-flattener designed according to Section 4 above, and the plate shifted a distance 0·5 inch towards the mirror when the field-flattener is in use, so as to balance the coma. Thus the fifth-order aberration-function of the field-flattened system is given by (57) and (66); that of the system without the field-flattener by (10).

As already remarked, it is possible to calculate the diffraction images of the two systems by means of these two fifth-order aberration-functions. But the computation is very laborious, on account of the numerical integrations involved. A good idea of the quality of the off-axis images can be obtained, with comparatively little labour, by plotting in the (X, Y)-plane the deviations $(\delta X, \delta Y)$ of the set of rays through the corner points

$$u, v = 0, \pm 0{\cdot}1, \pm 0{\cdot}2, \ldots ; \quad u^2 + v^2 \leqslant 1$$

of a lattice of small squares filling the entry-pupil. This gives a "spot diagram" in which the spot density distribution represents fairly accurately the intensity distribution in the geometrical images.

Fig. 8 shows the spot diagrams of the field-flattened system, in C, F and h-light, at four selected points of the field, namely $V = 0$, $V = \frac{1}{2}$, $V = \frac{1}{2}\sqrt{3}$ and $V = 1$. The second of these is one of the points of maximum coma and the third is the point at which the coma is in the act of changing its sign. Fig. 9 shows the corresponding spot diagrams for the system without the field-flattener.

From Figs. 7, 8 and 9 it appears that the addition of the field-flattener results in a marked increase in the colour-error of the sytem, and since colour-error

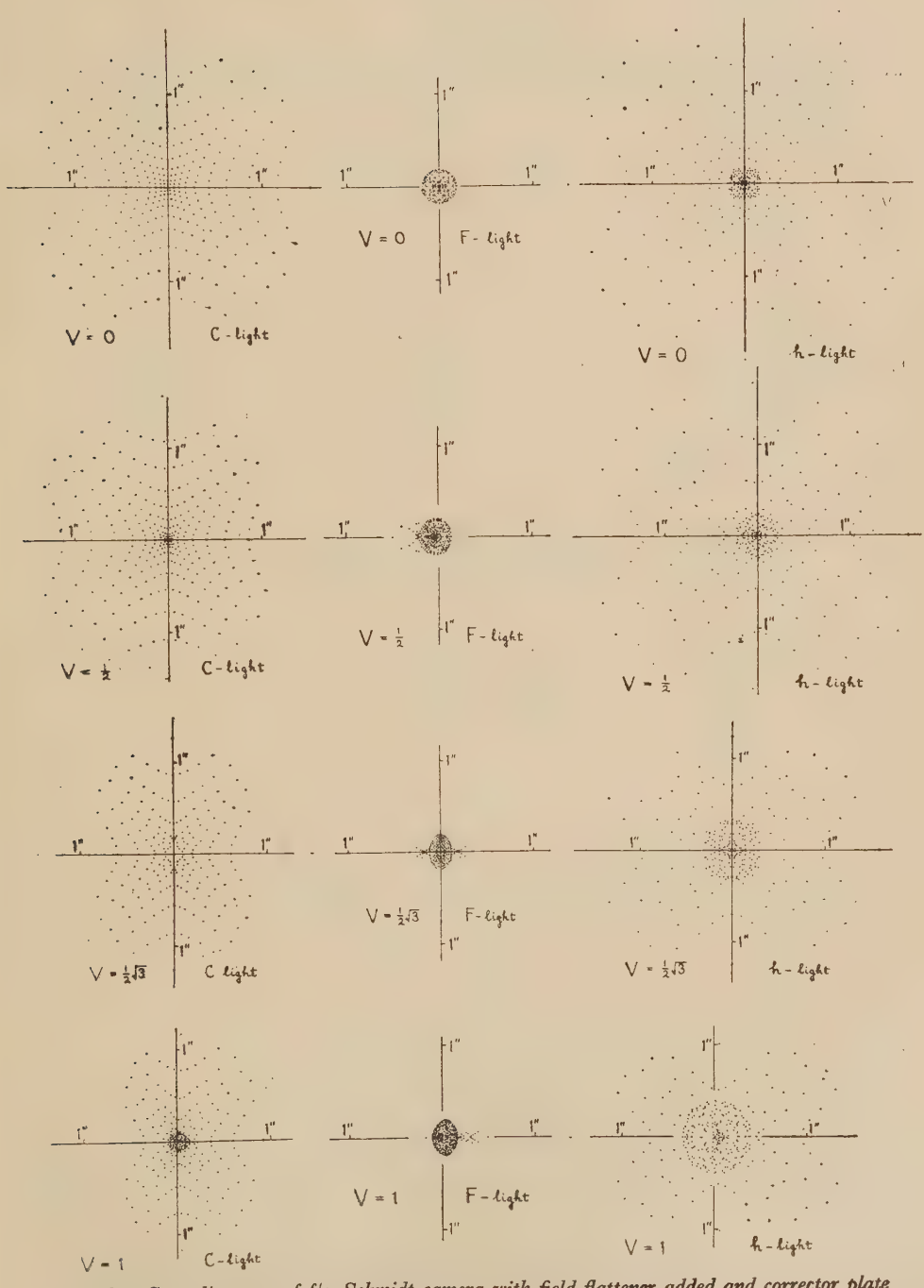

FIG. 8.—*Spot diagrams of f/3 Schmidt camera with field-flattener added and corrector plate moved to balance coma, at selected points of a 5° field. Plate-profile of type* $r^4 - \frac{4}{3}r^2$.

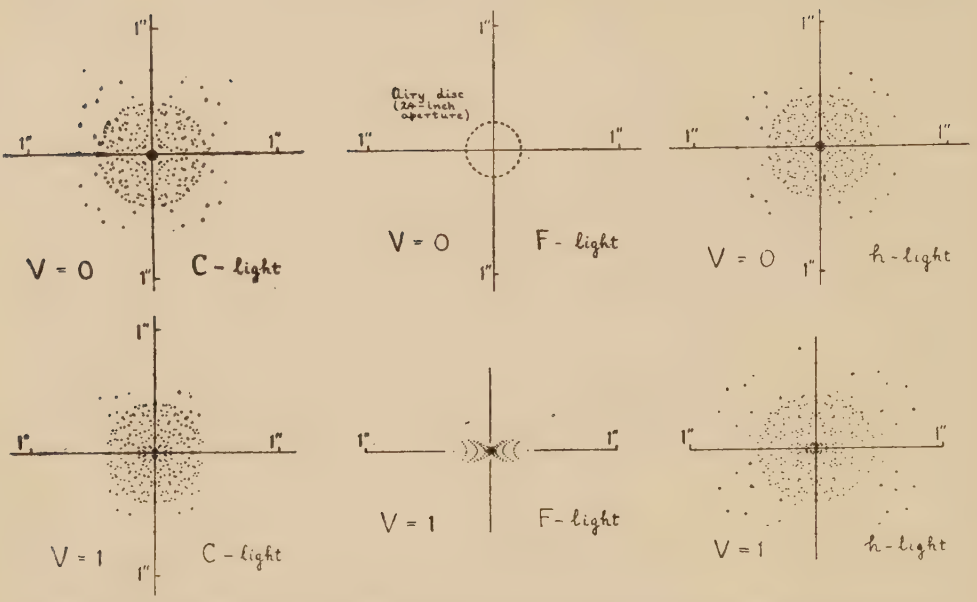

FIG. 9.—*Spot diagrams of f/3 Schmidt camera, axially stigmatic in F-light and with plate profile of type* $r^4 - \frac{4}{3}r^2$, *at the centre* ($V=0$) *and edge* ($V=1$) *of a* 5° *field.*

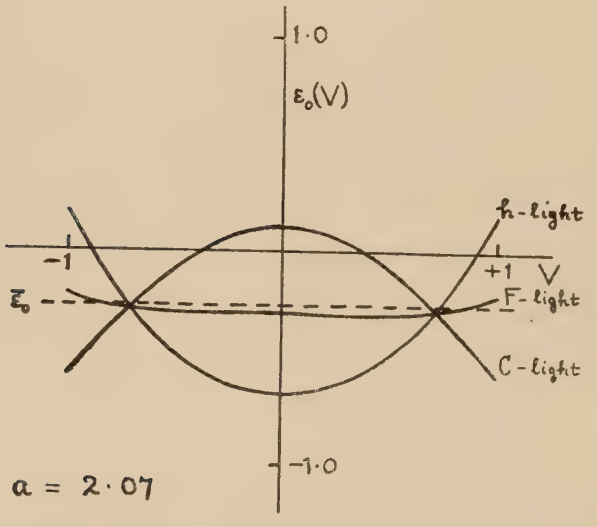

FIG. 10.—*Surfaces of best focus in the field-flattened system of Fig.* 11. *The dotted line represents the adopted focal surface.*

is the dominant aberration in a Schmidt camera of this focal ratio and field-size, the increase in colour-error carries with it a considerable loss in theoretical resolution.

An increase in the value of a, i. e. in the strength of the "central bulge" on the corrector plate, would improve this colour-error and could also be expected to reduce the monochromatic aberrations.* It therefore seems worth while to consider what improvement in performance can be obtained by designing the

* Compare (3), p. 756.

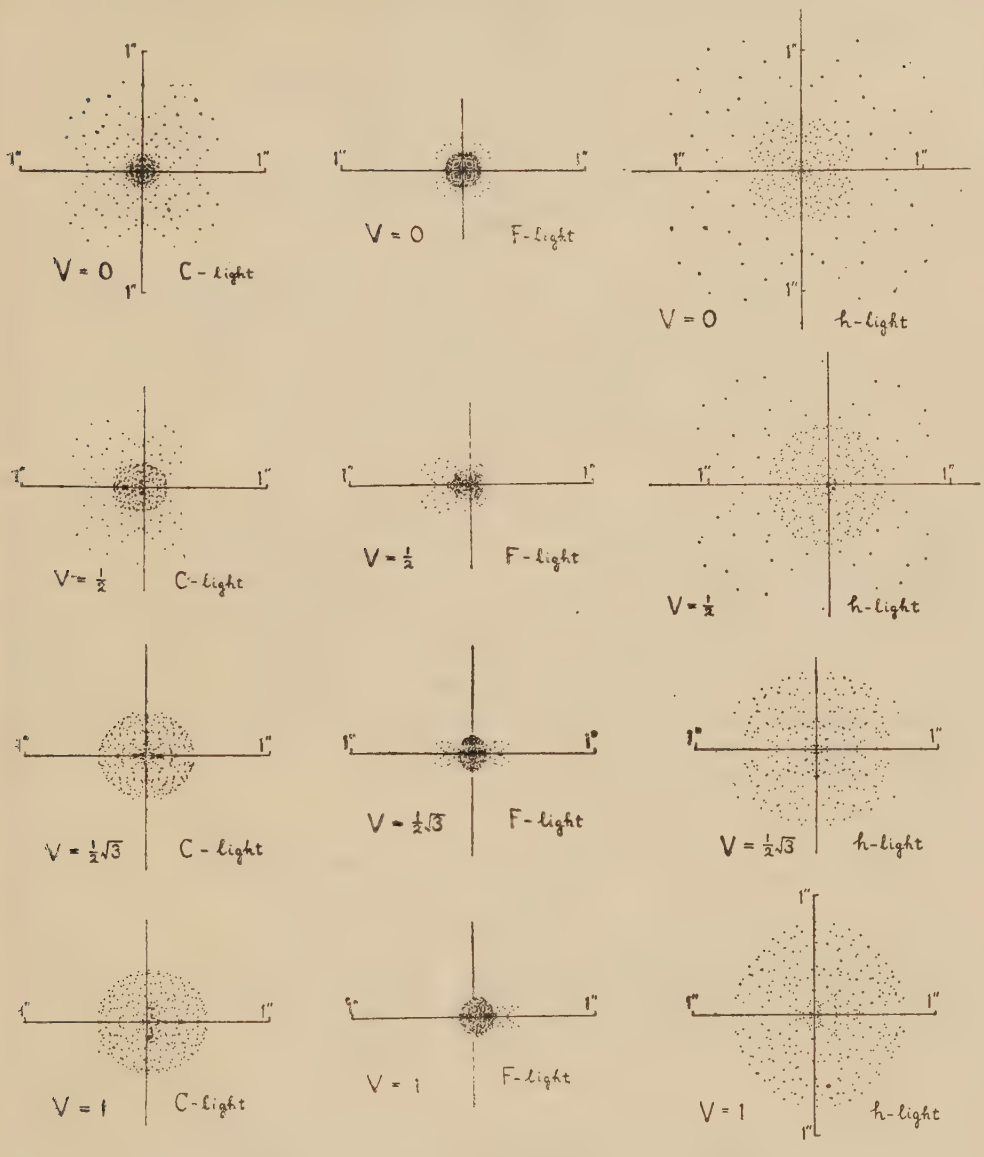

FIG. 11.—*Spot diagrams of f/3 field-flattened Schmidt camera, with plate-profile of type* $r^4 - 2{\cdot}07r^2$, *at selected points of a* 5° *field.*

system *ab initio* as a field-flattened Schmidt, instead of taking a plain Schmidt, adding a field-flattener, and moving the corrector plate to balance the coma.

In order to improve the chromatism, we first increase a until the C and h curves in Fig. 7 have been brought as close together as possible.* To do this, we set

$$\epsilon_C(1) - \epsilon_h(1) = \epsilon_h(0) - \epsilon_C(0), \tag{69}$$

* Owing to the dominance of colour-error over the monochromatic aberrations at this focal ratio, the further refinement of dropping axial F-stigmatism in the original Schmidt would not lead to any considerable improvement in the optical images.

where $\epsilon_C(V)$, $\epsilon_h(V)$ are the values of $\epsilon_0(V)$ in C and h-light respectively, and treat this as an equation for a. By (62) it can be written

$$4(a-\tfrac{4}{3})(k_{nn_{h\mu}}-k_{nn_{C\mu}})=\frac{8c^2}{n_1^2}\cdot\frac{7}{3}\cdot 4n_1(k_{n_1n_{1h\mu}}-k_{n_1n_{1C\mu}}),$$

i. e.

$$(a-\tfrac{4}{3})\frac{n_h-n_C}{n-1}=\frac{8c^2}{n_1}\cdot\frac{7}{3}\cdot\frac{(n_1)_h-(n_1)_C}{n_1-1}. \tag{70}$$

Here

$$c=\tfrac{1}{4},\ n_h-n_C=0{\cdot}0177,\ (n_1)_h-(n_1)_C=0{\cdot}0171,\ n-1=0{\cdot}52,\ n_1-1=0{\cdot}52.$$

Thus

$$a-\tfrac{4}{3}=\frac{1}{2n_1}\cdot\frac{7}{3}\cdot\frac{171}{177}=0{\cdot}74,$$

$$a=2{\cdot}07. \tag{71}$$

Fig. 10 shows the graphs of $\epsilon_0(V)$ in C, F and h-light for a system with $a=2{\cdot}07$. The chromatic variation of best focus is seen to be reduced by a factor 3. A trace of residual field-curvature in F-light can now be noticed, but it is too small in comparison with the colour-error to be worth removing. It will be seen that the surfaces of best focus in C and h-light are approximately symmetrical about the focal plane $\epsilon=\bar{\epsilon}_0=-0{\cdot}26$. Finally, spot diagrams for the new system are shown in Fig. 11, from which it appears that its geometrical aberrations are not much larger than those of the classical Schmidt camera.

The Observatory,
Cambridge:
1949 *August* 15.

References

(1) F. E. Ross, *Ap. J.*, **92**, 400–407, 1940.
(2) E. H. Linfoot, *M.N.*, **109**, 279–297, 1949.
(3) E. H. Linfoot and E. Wolf, *J.O.S.A.*, **39**, 752–756, 1949.

STELLAR PHOTOMETRY WITH A ROTATING SECTOR

Donald A. MacRae

(Communicated by C. P. Olivier)
(Received 1949 August 26)

Summary

Tests of a rotating sector were carried out at the telescope under ordinary conditions of astronomical photometry. It was found, in agreement with other investigations made in the laboratory, that, if the frequency of interruption is sufficiently high, interposing a sector is photographically equivalent to reducing the light intensity by the amount corresponding to the geometry of the sector. These results, taken in conjunction with the other properties of the sector, serve to emphasize its usefulness in establishing a stellar magnitude scale.

Introduction.—In a recent communication, Cousins and Stoy * have described some experiments, carried out in the laboratory, on the effectiveness of a rotating sector for reducing the intensity of a light source. As these authors point out, the rotating sector has not been used in the past for photometric purposes because of the uncertainty as to whether its action is equivalent to a reduction of the light intensity or a reduction in the exposure time. If the latter is the case the sector is of little use, for then the reciprocity failure of the photographic plate at astronomical intensities will give rise to a larger photographic effect than that produced by a continuous exposure at lower intensity, even though the total energy falling on the plate remains the same. The behaviour of a photographic plate exposed behind a rotating sector has been carefully investigated by Webb † at the Eastman Kodak Research Laboratories. He showed that, if the frequency of the interruption is sufficiently high, using a rotating sector is photographically equivalent to reducing the intensity of the light source. Cousins and Stoy, calibrating their light intensities in the laboratory by means of the inverse square law, confirmed Webb's results by showing that, with a sector giving 100 and more interruptions per second, the sector constant does not differ sensibly from that calculated from geometrical measurements.

If these conclusions are applicable under astronomical conditions the rotating sector provides us with a singularly effective means of setting up a magnitude scale. Unlike a diaphragm or a grating, placed over the telescope objective, the sector neither changes the character of the stellar image nor gives an optical path which is not representative of the beam as a whole. It is unique in being equally effective in light of all wave-lengths, although " neutral " density filters consisting of thin metallic films fulfil this condition very closely over a wide range of wave-length.

The experiments described in this paper were undertaken in the winter of 1941 and the following spring. The purpose was to test Webb's results in the telescope with in-focus stellar images under average conditions of seeing, and thus to determine whether or not the argument in favour of the sector could

* A. W. J. Cousins and R. H. Stoy, *M.N.*, **106**, 286, 1947.
† J. H. Webb, *J. Opt. Soc. Amer.*, **23**, 157, 1933.

be carried one step further. Observing conditions generally were chosen to match those met with in common practice. Experiments of this type are prerequisite to a critical evaluation of the sector as a photometric instrument of general application.

Procedure at the telescope.—The tests were conducted with the 10¼-inch Ross-Fecker camera at the Cook Observatory of the University of Pennsylvania. The sector and the mounting which held it immediately in front of the plate were constructed by Mr Lewis P. Tabor. The sector was cut from a disk of coated aluminium ten inches in diameter. There were eight openings, each about eleven degrees in width. When the electric motor was running at its normal speed the beam was interrupted 100 times a minute. Measurements of the cut-away portion of the disk established that the geometrical constant of the sector was 1·53 magnitudes with a mean error of three or four hundredths of a magnitude. The mounting made it possible to cover either the north half or the south half of the plate with the sector disk. The plates used were Eastman Spectroscopic I–G, exposed behind a Wratten number 15 (G) filter, so that the colour system approached that of the International Photovisual system. Exposures were centred on the star BD +84° 383. The plates, which were cut to four by five inches, covered an area of six and a half square degrees. Thus only a small central portion of the normal 20 by 24 inch field was used. The length of the exposures was generally sixty minutes, and stars of the eleventh magnitude produced well-blackened images. The plate limit was about magnitude 12·5. Tray development in D–19 was used.

First, two plates were taken of the field without the sector. Then a pair of plates were taken with the sector running at its normal speed, one plate with the sector covering the north half of the field, and the other with the sector covering the south half. These sector plates provided the means of reducing the measures on the field plates to magnitudes, with an arbitrary zero point. Additional sector plates were obtained to test the effect of reducing the speed of rotation of the sector, and the effect of sky fog. Finally, on one plate the sector was not used but the exposure on one-half of the field was composed of four four-minute exposures, while the other half received a continuous exposure of 66 minutes. This was done in order to obtain a measure of the maximum reciprocity failure of the plates under the conditions of the other experiments.

Results.—Measurements were carried out by means of a graduated scale of stellar images. All the stars in the field which are listed in " Magnitudes and colors of stars north of +80° " by Seares, Ross and Joyner * were measured in this way.

The reduction of the measures was carried out by a method essentially equivalent to the classical methods used with gratings and neutral screens, and adequately described in the literature.† The characteristic curve which is obtained relates scale readings on the field stars to magnitudes, with an arbitrary zero point. A fundamental assumption made in the reduction is that the effect of the sector was to diminish the light by 1·53 magnitudes. The test of this assumption is to compare the resultant characteristic curve with the one obtained by plotting the catalogue magnitudes against the scale readings of the field stars.

* F. H. Seares, F. E. Ross and M. C. Joyner, *Papers of the Mt. Wilson Obs.*, **6** (*Carnegie Inst. Pub.*, 532), 1941.

† For example : O. Heckmann and H. Kienle, *Veröff. Univ.-Sternw. zu Göttingen*, No. 52, 1937.

The catalogue magnitudes were first reduced to the colour system of the Cook magnitudes. The colour coefficient of the Cook lens when used with this type of plate and filter was found to be 0·13 mag. per magnitude of colour index, in the sense that the Cook measures of a red star are too bright. The Mt. Wilson magnitudes were adjusted accordingly, and then plotted against the scale readings. These points are represented in Fig. 1. The full-drawn curve in Fig. 1 is the characteristic curve determined from the sector plates on the assumption mentioned above, that the proper sector constant was the geometrical one.

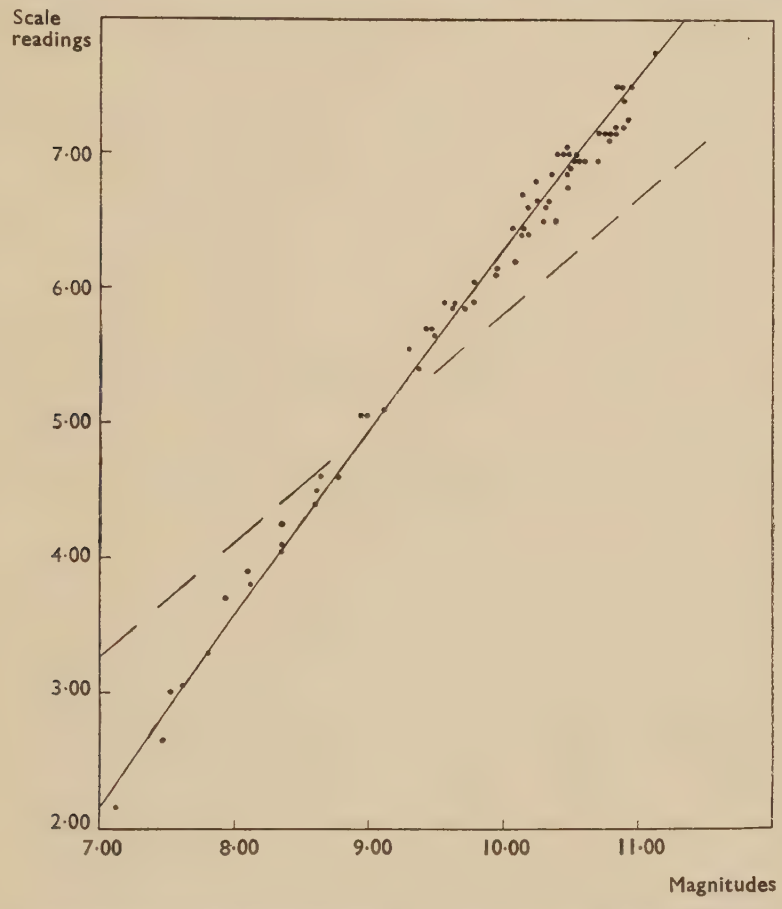

FIG. 1.—*Comparison of characteristic curves. The solid curve was determined from the sector plates; its zero point was adjusted to give the best fit to the points. The points were plotted from the known magnitudes of the stars. The dashed curve cannot be made to fit the points ; its slope was determined from plates taken under conditions of full reciprocity failure.*

The agreement of points and curve in Fig. 1 is remarkably good. Considering the full range of four magnitudes, there is no evidence that the adopted sector constant is in error. The fact that the points fall above the line in the middle region can be ascribed to the process of reduction in which the calibration curve (relating images with and without the sector) was drawn in as a straight line. If a slight curvature had been permitted in this calibration curve the points could have been more nearly represented. A larger number of sector plates would be required, however, in order to establish the existence of such a curvature.

Fig. 2 presents the same results in a different way. Here the differences between the Cook magnitudes and the adjusted Mt. Wilson magnitudes are plotted against magnitude. The dashed curve shows the run of the differences. The overall trend is not systematic and the greatest average deviation is 0·06 mag. The indications are that the sector constant determined geometrically is correct to within 0·04 mag.

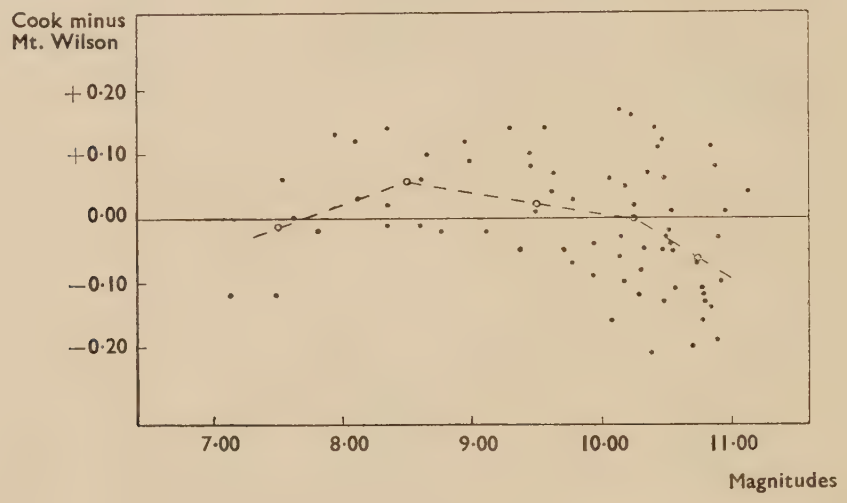

FIG. 2.—*Comparison of magnitude scales. The Cook magnitudes were read from the characteristic curve determined from the sector plates (solid curve of Fig. 1). The Mt. Wilson magnitudes were reduced to the colour system of the Cook telescope.*

The plate with four four-minute exposures on one-half was taken to simulate a sector plate for which the frequency of the interruption is extremely low; sixteen is the same fraction of 66 as the cut-away part of the sector was of the full circle. The four exposures were equally distributed over the 66-minute period to allow for possible variations in sky transparency. An exposure of this type is manifestly one in which the exposure time and not the light intensity is reduced. The full effect of reciprocity failure should therefore be expected. When the measures were reduced the large deviation in the magnitude scale was inescapable. In Fig. 1 the dashed curve is the characteristic curve which was obtained.

These last results can be expressed in another way by calculating what the sector constant would have had to be in order to reproduce the full curve in Fig. 1. This "effective sector constant" was 0·96 mag., or more than half a magnitude less than the calculated value. The mean error of this figure was of the order of 0·07 mag.

The plates taken with the sector running at slower speeds, so as to provide 2700 and 100 flashes per hour, were reduced in the same manner. The effective sector constants were 1·42 mag. and 1·35 mag. respectively. The difference between the second figure and 1·53 is significantly larger than its mean error.

It was difficult to obtain badly fogged plates because of the large focal ratio of the camera, *f*/8. On most plates the differential fog density was barely detectable. On one pair of normal sector plates, which showed fog densities of 0·19 and 0·16 on the uncovered and covered halves respectively, the effective sector constant was 1·48 mag. This suggests that the background fog is not

serious. Images close to the limit, however, were not used. Miller and Hynek *, with plates taken under more extreme fog conditions, found that the influence of fog on their filter and sector constants was small.

If observing conditions are such that some sky fog is unavoidable, it seems probable that its effect on the sector constant can be reduced to negligible proportions if the sector is constructed of a translucent material. The experiments at the Cook Observatory had to be discontinued before such a sector could be tested. A translucent sector would effectively interrupt the light from the star, while at the same time it would allow the light from the sky to continue to act on the plate at a large fraction of its full intensity. Thus all parts of the plate would be nearly equally affected and the resultant magnitude scale should be virtually free of errors arising from sky fog.

It is interesting to compare these results with those of Webb† with respect to light intensity and the critical frequency of flash. He came to the conclusion that the critical frequency measures the rate at which some kind of discrete process takes place in the photographic grain during reception of the light, and that this corresponds in order of magnitude to the rate of incidence of light quanta. Taking into account the area of the objective and the size of the stellar image, we find that the average number of quanta per square centimetre per second in the tests decribed here are about the same as in Webb's experiments. The two investigations agree in placing the critical frequency of flash at about one per second. The results of the present tests, in respect to the magnitude of the reciprocity failure, agree with the published curve for the I–G emulsion.‡

Conclusion.—The rotating sector has been shown, within the limits of the experimental accuracy, to be a reliable instrument for reducing the intensity of a star's light. The effect of reciprocity failure is avoided if the frequency of interruption is made sufficiently high. The critical frequency seems to be about the same as the frequency of incidence of light quanta on the photographic grain; under the conditions of stellar photometry described here it is about one flash per second. A factor of safety of 100 or 1000 is probably advisable. This device should find a wider use than heretofore in general programmes of astronomical photometry because of its peculiar advantages, mentioned in the first section.

Flower Astronomical Observatory,
Philadelphia, Pennsylvania :
1949 *August* 19.

* F. D. Miller and J. A. Hynek, *Cont. Perkins Observatory*, No. 13, 1939.

† J. H. Webb, *J. Opt. Soc. Amer.*, **23**, 316, 1933.

‡ *Photographic plates for use in spectroscopy and astronomy*, Eastman Kodak Company, Third Ed., p. 12, 1937.

ON THE EFFECT OF RADIATION PRESSURE IN THE ATMOSPHERES OF EARLY-TYPE STARS

Anne B. Underhill *

(Received 1949 April 11)

Summary

The effects of radiation pressure on the stability of the atmospheres of the high-temperature stars are discussed and it is shown that these effects are not negligible for stars with $T_e > 20{,}000$ deg. An expression for the adiabatic gradient, $d \log T / d \log p_g$, in a mixture of radiation and an ionizable gas in thermodynamic equilibrium is obtained. A recent method, proposed by Unsöld, of obtaining the adiabatic gradient from entropy considerations is extended to include the effect of the radiation field. It is shown that stars of high effective temperature and low surface gravity probably do not have stable atmospheres.

In recent years model atmospheres have been extensively studied by B. Strömgren (**1**), but if his methods are to be adapted to the study of high-temperature atmospheres the effects of radiation pressure must be closely considered. Since we postulate that the atmosphere is in mechanical equilibrium, we have

$$\frac{d(p_r + p_g)}{dz} = -g\rho, \tag{1 a}$$

where p_g is the gas pressure at any level, z, in the atmosphere, p_r is the radiation pressure, ρ is the density at this level and g is the surface gravity. In the investigations (**1**) of model atmospheres referred to above, the term in equation (1 *a*) referring to the radiation pressure is neglected, and the equation of mechanical equilibrium is written as

$$\frac{dp_g}{dz} = -g\rho. \tag{1 b}$$

This equation is integrated to give the structure of the stellar atmosphere. This procedure is satisfactory so long as the effective temperature is low, but if $T_e > 20{,}000$ deg. the term in the radiation pressure can no longer be neglected. Rudkjöbing, in his discussion (**2**) of model atmospheres for the B stars, recognizes this fact, and following Unsöld (**3**) allows for the effect of radiation pressure by defining an effective surface gravity such that

$$g_{\text{eff}} = g - g_r, \tag{2}$$

where g_r is an acceleration, assumed to be constant with depth, which gives the effect of radiation pressure. We have

$$g_r = -\frac{1}{\rho}\frac{dp_r}{dz}. \tag{3}$$

Rudkjöbing then determines the structure of his model atmospheres using the equation of mechanical equilibrium

$$\frac{dp_g}{dz} = -g_{\text{eff}}\rho. \tag{4}$$

However, g_r is not quite constant with depth even in the rather favourable models (from the point of view of small effect of radiation pressure) calculated by

* Fellow of the National Research Council (U.S.A.).

Rudkjöbing, and one can readily estimate that this variance will become more serious in models at higher T_e and lower values of g. A more pliant method of dealing with the effects of radiation pressure than this is required if model atmospheres are to be calculated for the more luminous high-temperature stars. Such a method will be developed here.*

It may readily be shown from fundamental principles that if $I_\nu(\vartheta, z)$ is the intensity of radiation of frequency ν in a beam inclined at an angle ϑ to the normal at any level z in the atmosphere, the pressure due to the radiation in frequency ν is

$$p_r(\nu) = \frac{2\pi}{c}\int_0^\pi I_\nu(\vartheta, z)\cos^2\vartheta\sin\vartheta\,d\vartheta. \tag{5}$$

The total pressure due to the radiation field in all frequencies is

$$p_r = \int_0^\infty p_r(\nu)d\nu, \tag{6}$$

hence

$$\frac{dp_r}{dz} = \frac{2\pi}{c}\int_0^\infty d\nu\int_0^\pi \frac{dI_\nu(\vartheta, z)}{dz}\cos^2\vartheta\sin\vartheta\,d\vartheta. \tag{7}$$

Now the monochromatic equation of transfer in the atmosphere of a high-temperature star in which the opacity is given by electron scattering as well as by continuous absorption is

$$\cos\vartheta\frac{dI_\nu(\vartheta, z)}{dz} = -(\kappa_\nu+\sigma)\rho I_\nu(\vartheta, z) + \tfrac{1}{2}\sigma\rho\int_0^\pi I_\nu(\vartheta, z)\sin\vartheta\,d\vartheta + j_\nu(z), \tag{8}$$

where κ_ν is the mass absorption coefficient and σ is the mass scattering coefficient which to a good approximation is independent of frequency. The last term gives the emission which is a function of depth in the atmosphere only. If we substitute expression (8) for $\cos\vartheta\,dI_\nu/dz$ in equation (7) and perform the integration over ϑ we obtain

$$\frac{dp_r}{dz} = -\frac{2\pi}{c}\int_0^\infty(\kappa_\nu+\sigma)\rho\,d\nu\int_0^\pi I_\nu(\vartheta, z)\sin\vartheta\cos\vartheta\,d\vartheta. \tag{9}$$

Since the monochromatic flux is defined by

$$F_\nu = 2\int_0^\pi I_\nu(\vartheta, z)\sin\vartheta\cos\vartheta\,d\vartheta,$$

we can write

$$\frac{dp_r}{dz} = -\frac{\pi}{c}\int_0^\infty(\kappa_\nu+\sigma)\rho F_\nu\,d\nu. \tag{10}$$

We shall now introduce a mean absorption coefficient $\bar{\kappa}$ such that at any depth z

$$\kappa_\nu = \bar{\kappa}(1+\delta_\nu), \tag{11}$$

where δ_ν is a quantity giving the departure from greyness. Then we obtain

$$\frac{dp_r}{dz} = -\frac{\pi}{c}(\bar{\kappa}+\sigma)\rho\int_0^\infty F_\nu d\nu - \frac{\pi}{c}\bar{\kappa}\rho\int_0^\infty\delta_\nu F_\nu d\nu. \tag{12}$$

If $\bar{\kappa}$ is a Chandrasekhar-type mean absorption coefficient (**4, 5**), we have explicitly

$$\int_0^\infty\delta_\nu F_\nu^{(1)}d\nu = 0,$$

* Some parts of the ideas to be developed in this paper are contained with rather different emphasis in an early paper by E. A. Milne, *M.N.*, **85**, 768, 1925.

where $F_\nu^{(1)}$ is the grey-body monochromatic flux. Consequently we can write

$$\frac{dp_r}{dz} = -\frac{\pi}{c}(\bar{\kappa}+\sigma)\rho\int_0^\infty F_\nu d\nu - \frac{\pi}{c}\bar{\kappa}\rho\int_0^\infty \delta_\nu(F_\nu - F_\nu^{(1)})d\nu. \tag{13}$$

The last term will vanish for a grey atmosphere, i.e. when $\bar{\kappa}\ll\sigma$, for σ is independent of frequency. Hence for a grey atmosphere we have

$$\frac{dp_r}{dz} = -\frac{\pi}{c}(\bar{\kappa}+\sigma)\rho\int_0^\infty F_\nu d\nu. \tag{14}$$

The definition of effective temperature gives

$$\pi\int_0^\infty F_\nu d\nu = \sigma_R T_e^4, \tag{15}$$

where σ_R is the radiation constant, $5{\cdot}67\times10^{-5}$ ergs/cm.^{2}/sec./deg.^{4}. Hence we find for a grey atmosphere that

$$\frac{dp_r}{dz} = -\frac{\sigma_R}{c}(\bar{\kappa}+\sigma)\rho\, T_e^4. \tag{16}$$

The equation of mechanical equilibrium, upon introducing the optical depth

$$d\tau = -(\bar{\kappa}+\sigma)\rho\, dz, \tag{17}$$

becomes

$$\frac{dp_g}{d\tau} = \frac{g}{\bar{\kappa}+\sigma} - \frac{\sigma_R}{c}T_e^4 \tag{18}$$

for a grey atmosphere. For a non-grey atmosphere the term

$$\frac{\pi}{c}\frac{\bar{\kappa}}{\bar{\kappa}+\sigma}\int_0^\infty \delta_\nu(F_\nu - F_\nu^{(1)})d\nu$$

must be added to the right side of equation (18).

In equation (18) we have a form of the equation of mechanical equilibrium which is convenient for computing model atmospheres of high effective temperature. In the outermost layers of any high-temperature atmosphere it is a good approximation, for there $\sigma>\bar{\kappa}$ and the atmosphere is "grey" to a good approximation. At deeper layers $\bar{\kappa}>\sigma$, and then the additional term may play a role. The effect of this term can only be evaluated by numerical investigations, and is closely connected with the difference between the effective temperature found for a model atmosphere by the evaluation of equation (15) numerically, and that obtained from any given theoretical temperature law and an assumed boundary temperature.*

We shall accept the grey-body approximation for treating the effects of radiation pressure on the mechanical equilibrium of a stellar atmosphere and shall examine further the meaning of equation (18). However, before doing this we note from equation (3) that

$$g_r = \frac{\sigma_R}{c}(\bar{\kappa}+\sigma)T_e^4. \tag{19}$$

Since in any atmosphere $(\bar{\kappa}+\sigma)$ is not usually constant with depth, we see that we cannot expect g_r to be constant with depth. The use of equation (18) obviates any necessity of defining an effective acceleration of gravity and then hoping that it will be constant with depth.

* See A. B. Underhill, "On Model Atmospheres for the High-Temperature Stars", *Publ. Copenhagen Obs.*, in press, for a discussion of these points.

Table I gives the value of $dp_r/d\tau$ for a number of values of T_e. Also given is $dp_g/d\tau$ for two typical values of $g/(\bar{\kappa}+\sigma)$. We see that the gas pressure gradient can become quite small, or even negative. Such behaviour suggests that the

TABLE I

The Effect of Radiation Pressure on the Pressure Gradient

T_e (deg.)	$\frac{dp_r}{d\tau}=\frac{\sigma_R}{c}T_e^4$	$dp_g/d\tau$	
		$g/(\bar{\kappa}+\sigma)=10^4$	$g/(\bar{\kappa}+\sigma)=10^3$
5000	1·18	$1\cdot0\times10^4$	$1\cdot0\times10^3$
10,000	18·9	$1\cdot0\times10^4$	980
20,000	302	$9\cdot7\times10^3$	700
25,000	737	$9\cdot3\times10^3$	260
30,000	$1\cdot53\times10^3$	$8\cdot5\times10^3$	−530
35,000	$2\cdot84\times10^3$	$7\cdot2\times10^3$	$-1\cdot84\times10^3$
40,000	$4\cdot84\times10^3$	$5\cdot2\times10^3$	$-3\cdot84\times10^3$
45,000	$7\cdot75\times10^3$	$2\cdot2\times10^3$	$-6\cdot75\times10^3$
50,000	$1\cdot18\times10^4$	$-1\cdot8\times10^3$	$-1\cdot08\times10^4$

atmosphere will become convectively unstable. K. Schwarzschild has shown that if the existing temperature and pressure gradients in an atmosphere are such that

$$\left|\frac{d\log T}{d\log p_g}\right|>\left|\left(\frac{d\log T}{d\log p_g}\right)_{\text{adiabatic}}\right|, \tag{20}$$

convection will set in. At any level in an atmosphere

$$\frac{d\log T}{d\log p_g}=\frac{p_g}{T}\left(\frac{dp_g}{d\tau}\right)^{-1}\frac{dT}{d\tau}, \tag{21}$$

hence unless $dT/d\tau$ becomes very small as $dp_g/d\tau$ becomes small, we may expect $d\log T/d\log p_g$ to become quite large and convection to occur.

The question of the value of the adiabatic gradient at any level in the atmosphere requires some consideration. Chandrasekhar (**6**) has considered the case of an enclosure containing a perfect gas and radiation in thermodynamic equilibrium. He shows that for an adiabatic change

$$(12p_r+\frac{1}{\gamma-1}p_g)\frac{dT}{T}+(4p_r+p_g)\frac{dV}{V}=0, \tag{22}$$

where γ is the ratio of the specific heats of a perfect gas. For a perfect gas

$$\frac{dV}{V}=\frac{dT}{T}-\frac{dp_g}{p_g}, \tag{23}$$

and consequently it follows that for radiation and a perfect gas in an enclosure,

$$\left(\frac{d\log T}{d\log p_g}\right)_{\text{adiabatic}}=\frac{4p_r/p_g+1}{16p_r/p_g+\frac{5}{2}}. \tag{24}$$

Unsöld (**7**) considers the adiabatic gradient in a gas, one component of which is being ionized. He ignores possible contributions from the radiation field. His discussion can readily be modified to include the effect of the radiation field, and thus to yield an expression suitable to our purpose.

We shall consider a volume V containing radiation and N atoms or ions in thermodynamic equilibrium at temperature T. The various elements are present

in relative abundances ν_i, $(\Sigma\nu_i=1)$, and each is ionized to a degree x_i. Only element 1 is changing its degree of ionization, hence all

$$dx_i=0, \qquad i\neq 1.$$

Since our gas is a perfect gas,

$$p_g V=NkT(1+\bar{x}), \tag{25}$$

where we have written

$$\bar{x}=\Sigma\nu_i x_i.$$

Logarithmic differentiation of equation (25) gives

$$\frac{dp_g}{p_g}+\frac{dV}{V}=\frac{dT}{T}+\frac{\nu_1 dx_1}{1+\bar{x}}. \tag{26}$$

From Saha's equation we have

$$\frac{x_1}{1-x_1}\cdot\frac{\bar{x}}{1+\bar{x}}p_g=\text{const}\ T^{\frac{5}{2}}e^{-\chi_1/kT}, \tag{27}$$

since

$$p_e=\frac{\bar{x}}{1+\bar{x}}p_g,$$

and when we differentiate equation (27) logarithmically we obtain

$$dx_1\left[\frac{1}{x_1(1-x_1)}+\frac{\nu_1}{\bar{x}(1+\bar{x})}\right]=\left[\tfrac{5}{2}+\frac{\chi_1}{kT}\right]\frac{dT}{T}-\frac{dp_g}{p_g}. \tag{28}$$

We now substitute the value of dx_1 from equation (28) in equation (26) with the result that

$$\frac{dV}{V}=(A+1)\frac{dT}{T}-(B+1)\frac{dp_g}{p_g}, \tag{29}$$

where

$$B=\frac{\nu_1}{\dfrac{1+\bar{x}}{x_1(1-x_1)}+\dfrac{\nu_1}{\bar{x}}} \tag{30}$$

and

$$A=B\left(\tfrac{5}{2}+\frac{\chi_1}{kT}\right). \tag{31}$$

The internal energy of our system is the radiation energy in the volume plus the kinetic and potential energy of the particles. We have

$$U=VaT^4+\tfrac{3}{2}NkT(1+\bar{x})+N\Sigma\nu_i x_i\chi_i, \tag{32}$$

since the energy density for radiation in thermodynamic equilibrium at temperature T is

$$u=aT^4,$$

where a is the Stefan-Boltzmann constant, $7{\cdot}55\times 10^{-15}$ ergs/cm.^{3}/deg.^{4}. For an adiabatic change

$$dU+PdV=0, \tag{33}$$

where P is the total pressure, p_r+p_g. Hence

$$-\frac{dV}{V}=\frac{dU}{PV}. \tag{34}$$

Now

$$\frac{dU}{PV}=\frac{dT[4VaT^3+\frac{3}{2}Nk(1+\bar{x})]+aT^4dV+\nu_1 dx_1 NkT[\frac{3}{2}+\chi_1/kT]}{(p_r+p_g)V}. \tag{35}$$

However,

$$p_r = \tfrac{1}{3}aT^4$$

for radiation in thermodynamic equilibrium in an enclosure, and

$$p_g = \frac{NkT}{V}(1+\bar{x}),$$

hence

$$(p_r+p_g)\frac{dU}{PV} = \frac{dT}{T}[12p_r+\tfrac{3}{2}p_g]+3p_r\frac{dV}{V}+\frac{\nu_1 dx_1}{1+\bar{x}}\left[\tfrac{3}{2}+\frac{\chi_1}{kT}\right]p_g.$$

When we substitute for $\nu_1 dx_1/(1+\bar{x})$ from equation (26) we find

$$(p_r+p_g)\frac{dU}{PV} = \frac{dT}{T}\left[12p_r-p_g\frac{\chi_1}{kT}\right]+\frac{dV}{V}\left[3p_r+p_g\left(\tfrac{3}{2}+\frac{\chi_1}{kT}\right)\right]+dp_g\left[\tfrac{3}{2}+\frac{\chi_1}{kT}\right].$$

Using equation (34) we obtain

$$\frac{dT}{T}\left[12p_r-p_g\frac{\chi_1}{kT}\right]+\frac{dV}{V}\left[4p_r+p_g\left(\tfrac{5}{2}+\frac{\chi_1}{kT}\right)\right]+dp_g\left[\tfrac{3}{2}+\frac{\chi_1}{kT}\right]=0. \quad (36)$$

We have an expression for dV/V in equation (29). The introduction of this expression into equation (36) gives

$$\left(\frac{d\log T}{d\log p_g}\right)_{\text{adiabatic}} = \frac{4(B+1)p_r/p_g+\left[1+B\left(\tfrac{5}{2}+\frac{\chi_1}{kT}\right)\right]}{4(A+4)p_r/p_g+\left[\tfrac{5}{2}+A\left(\tfrac{5}{2}+\frac{\chi_1}{kT}\right)\right]}. \quad (37)$$

Equation (37) gives the adiabatic gradient for a mixture of radiation and an ionizable gas in thermodynamic equilibrium at a temperature T. This expression reduces to that given by Unsöld when $p_r/p_g=0$, that is when the radiation field can be neglected, and to equation (24), which is valid for a mixture of radiation and a perfect non-ionizable gas, when $B=A=0$, that is when no component of appreciable abundance is being ionized. The effect of radiation pressure and of the possibility of ionization is to decrease the value of the adiabatic gradient from its value, 0·400, for a perfect gas in thermodynamic equilibrium. Since the effect of radiation pressure is to reduce the adiabatic gradient and also to increase the existing temperature-pressure gradient in the stellar atmosphere, *cf.* equation (21), we see that an atmosphere in which the radiation pressure is an appreciable fraction of the gas pressure may readily become unstable against convection.

The question of determining the adiabatic gradient, $d\log T/d\log p_g$, in a stellar atmosphere has recently been considered by Unsöld (**8**). He computes the entropy of a mixture of H, H^+, He, He^+, He^{++} and electrons in thermodynamic equilibrium as a function of $\log T$ and $\log p_g$, and from the locus of points which give $S=$ constant he obtains the adiabatic gradient by numerical differentiation. However, he does not consider the contribution of the radiation field to the entropy of the system. At high temperatures this contribution is not negligible and it should be included. Unsöld's method of attack has the advantage that the adiabatic gradient may be readily found even if several elements are being ionized at the same time, whereas the discussion leading to equation (37) cannot be carried through if more than one element is being ionized. The inclusion of the entropy from the radiation field will make Unsöld's discussion general, and will make it possible for one to obtain the adiabatic gradient under very general circumstances.

Following Unsöld we readily find that the entropy, $S_{\text{particles}}$, of N particles of H, H^+, He, He^+, He^{++} and the electrons set free in thermodynamic equilibrium at temperature T in volume V is given by

$$\left.\begin{aligned}\frac{S_{\text{particles}}}{kN} &= \tfrac{5}{2} + \ln\frac{(2\pi m_{\text{H}})^{\frac{3}{2}}k^{\frac{5}{2}}}{h^3} + \tfrac{3}{2}(\nu_1 \ln\mu_{\text{H}} + \nu_2 \ln\mu_{He}) + \tfrac{5}{2}\ln T - \ln p_e \\ &\quad + \ln[\nu_1 x_{\text{H}} + \nu_2(x_{He\,\text{I}} + 2x_{He\,\text{II}})] \\ &\quad - \nu_1\left[\ln\nu_1 + (1-x_{\text{H}})\ln\frac{1-x_{\text{H}}}{2} + x_{\text{H}}\ln x_{\text{H}}\right] \\ &\quad - \nu_2\left[\ln\nu_2 + (1-x_{He\,\text{I}}-x_{He\,\text{II}})\ln(1-x_{He\,\text{I}}-x_{He\,\text{II}})\right. \\ &\quad \left. + x_{He\,\text{I}}\ln\frac{x_{He\,\text{I}}}{2} + x_{He\,\text{II}}\ln x_{He\,\text{II}}\right] + [\nu_1 x_{\text{H}} + \nu_2(x_{He\,\text{I}} + 2x_{He\,\text{II}})] \\ &\quad \times\left[\tfrac{5}{2} + \ln\frac{(2\pi m_{\text{H}})^{\frac{3}{2}}k^{\frac{5}{2}}}{h^3} + \tfrac{3}{2}\ln\mu_{el} + \ln 2 + \tfrac{5}{2}\ln T - \ln p_e\right],\end{aligned}\right\} \quad (38)$$

where k is the Boltzmann constant, m_{H} is the mass of the hydrogen atom, ν_1 is the abundance by number of hydrogen, ν_2 is the abundance by number of helium, μ_{H}, μ_{He} and μ_{el} are the atomic weights of hydrogen, helium and the electron respectively, x_{H} is the degree of ionization of hydrogen, $x_{He\,\text{I}}$ gives the fraction of singly ionized helium atoms, and $x_{He\,\text{II}}$ gives the fraction of doubly ionized helium atoms. We find the degrees of ionization of hydrogen and helium as functions of the temperature and the electron pressure from the ionization equations

$$\log\frac{1-x_{\text{H}}}{x_{\text{H}}} = -13{\cdot}54\theta + \tfrac{5}{2}\log T - 0{\cdot}477 - \log p_e,$$

$$\log\frac{N_{He^+}}{N_{He^\circ}} = \log A = -24{\cdot}46\theta + \tfrac{5}{2}\log T + 0{\cdot}125 - \log p_e,$$

$$\log\frac{N_{He^{++}}}{N_{He^+}} = \log B = -54{\cdot}14\theta + \tfrac{5}{2}\log T - 0{\cdot}477 - \log p_e,$$

where $\theta = 5040/T$, and from the definitions

$$x_{He\,\text{I}} = \frac{N_{He^+}}{N_{He}} = \frac{A}{1+A(1+B)},$$

$$x_{He\,\text{II}} = \frac{N_{He^{++}}}{N_{He}} = \frac{AB}{1+A(1+B)}.$$

The gas pressure is related to the electron pressure by the relation

$$\frac{p_g}{p_e} = \frac{1+\nu_1 x_{\text{H}} + \nu_2(x_{He\,\text{I}} + 2x_{He\,\text{II}})}{\nu_1 x_{\text{H}} + \nu_2(x_{He\,\text{I}} + 2x_{He\,\text{II}})}. \quad (39)$$

Since the particles behave as a perfect gas we have the relation

$$p_g V = NkT[1 + \nu_1 x_{\text{H}} + \nu_2(x_{He\,\text{I}} + 2x_{He\,\text{II}})]. \quad (40)$$

If we denote the entropy of the radiation field in thermodynamic equilibrium in the volume V at temperature T by S_{rad}, the entropy of the system, radiation + particles, will be given by

$$S = S_{\text{particles}} + S_{\text{rad}},$$

since entropy is additive for two systems in thermodynamic equilibrium.

For radiation in thermodynamic equilibrium at temperature T in a volume V the entropy S is given by

$$S = \int \frac{dQ}{T},$$

where

$$dQ = dU + p_r dV,$$

U being the internal energy ($= aT^4V$). This leads on integration to the result

$$S_{\text{rad}} = \tfrac{4}{3} aT^3V, \tag{41}$$

where the constant of integration (which is immaterial in the application that follows) has been adjusted so as to make $S = 0$ for $T = 0$. By using equation (40) we obtain

$$\frac{S_{\text{rad}}}{Nk} = \frac{4}{3} \frac{aT^4[1 + \nu_1 x_{\text{H}} + \nu_2(x_{He\,\text{I}} + 2x_{He\,\text{II}})]}{p_g}. \tag{42}$$

The entropy of the system of particles and radiation in thermodynamic equilibrium at temperature T and gas pressure p_g is

$$\frac{S}{kN} = \frac{S_{\text{particles}}}{kN} + \frac{S_{\text{rad}}}{kN}, \tag{43}$$

where $S_{\text{particles}}/kN$ is found from equation (38) and S_{rad}/kN from equation (42). If this expression is plotted as a function of $\log T$ and $\log p_g$, the derivatives $d \log T / d \log p_g$ may be found numerically along the locus $S =$ constant in the manner proposed by Unsöld (**8**). In this way the adiabatic gradient may be found at any temperature and pressure. At moderately low gas pressures and for temperatures $> 10^4$ the effect of the entropy from the radiation field is to make the curves more vertical than in Unsöld's Figure 1. Such behaviour is in agreement with the effects of radiation pressure discussed above. The term (42) is negligible when T is less than about 20,000 deg., or if p_g is large and T is small. Consequently Unsöld's results should be modified to include the effect of the radiation field over most of his range of values.*

From equation (18) we can derive a minimum value of the surface gravity for which the gas pressure gradient in the outermost layers of a stellar atmosphere is positive. In order to have

$$\frac{dp_g}{d\tau} \geqslant 0, \tag{44}$$

we must have

$$g \geqslant (\bar{\kappa} + \sigma) \frac{\sigma_R}{c} T_e^4. \tag{45}$$

In the outermost layers of the stellar atmosphere the mean absorption coefficient per gram of star material, $\bar{\kappa}$, approaches zero. However, the coefficient of continuous scattering, σ, approaches a constant value of 0·397 if the atmosphere is pure hydrogen, or 0·315 if the atmosphere is 85 per cent by number *H* atoms and 15 per cent *He* atoms. (Here $\sigma = \sigma_e \times$ number of electrons/gram of star material, where σ_e is the scattering coefficient per electron.) If we use the value of 0·315 for the mass scattering coefficient, σ, we find the minimum g-values given in Table II. If the surface gravity is less than these values a stable atmosphere

* See A. B. Underhill, *Ann. d'Ap.* (in press).

cannot exist. These values are minimum values, for it may readily be shown that $\bar{\kappa}+\sigma>0{\cdot}315$ in most atmospheres.

TABLE II

Limiting Values of the Surface Gravity for a Mechanically Stable Atmosphere

T_e (deg.)	$g_{\min}$	T_e (deg.)	$g_{\min}$	T_e (deg.)	$g_{\min}$
5000	0·372	25,000	232	40,000	$1{\cdot}52\times10^3$
10,000	5·95	30,000	482	45,000	$2{\cdot}44\times10^3$
20,000	95·1	35,000	895	50,000	$3{\cdot}72\times10^3$

The data of Table II indicate that a mechanically stable fringe of atmosphere will not exist for stars of low surface gravity and high effective temperature. For these stars the total pressure gradient $d(p_g+p_r)/d\tau$ will exceed its equilibrium value $g/(\bar{\kappa}+\sigma)$ and either the atmosphere will expand as a whole, or more probably a large-scale turbulent motion will set in. This motion will be radial and will probably appear like prominence action. We should *not* expect the spectra from such stars to show the normal type of absorption lines which we know are formed in atmospheres in mechanical equilibrium. Consequently we conclude that all O stars having *normal* absorption-line spectra are stars of large surface gravity, that is they are main-sequence stars. This conclusion is in agreement with that drawn from the absence of emission lines in these stars (**4**). No decision about the absolute luminosities of the absorption-line O stars can be made from the observations at the moment as it has been so far impossible to find spectral criteria of differing absolute magnitude amongst the absorption-line O stars (**9**), and other indications of differing luminosities of these stars are inconclusive.

Beals (**10**) has estimated effective temperatures of the order of 50,000 deg. to 100,000 deg. for the Wolf-Rayet stars. For $T_e=100{,}000$ deg., $g_{\min}=5{\cdot}95\times10^4$, which is a value larger than the surface gravity estimated at present for main-sequence stars. The high effective temperatures of the Wolf-Rayet stars seem to be almost a sufficient guarantee that mechanical instability of their atmospheres occur. Should it be shown that the effective temperatures of the Wolf-Rayet stars are of the order of 50,000 deg. or less, we could conclude from the observed mechanical instability of their atmospheres that the Wolf-Rayet stars are stars of low surface gravity. At present there is no clear basis for such a conclusion.

University Observatory,
Copenhagen :
1949 *April* 7.

References

(1) B. Strömgren, *Festschrift für Elis Strömgren*, Copenhagen, *Publ. Copenhagen Obs.*, No. 127, 1940 ; and *Publ. Copenhagen Obs.*, No. 138, 1944. The latter publication contains a number of references to earlier work on model stellar atmospheres, by Milne, McCrea, Chandrasekhar, Russell, Unsöld and Pannekoek.
(2) M. Rudkjöbing, *Publ. Copenhagen Obs.*, No. 145, 1947.
(3) A. Unsöld, *Physik der Sternatmosphären*, Berlin, p. 141, 1938.
(4) A. B. Underhill, *Ap. J.* (in press).
(5) S. Chandrasekhar, *Ap. J.*, **101**, 328, 1945.
(6) S. Chandrasekhar, *An Introduction to the Study of Stellar Structure*, Chicago, p. 55 ff, 1939.
(7) A. Unsöld, *Physik der Sternatmosphären*, p. 380 ff, 1938.
(8) A. Unsöld, *Zs. f. Ap.*, **25**, 11, 1948.
(9) R. M. Petrie, *Publ. Dom. Ap. Obs.*, **7**, 321, 1947.
(10) C. S. Beals, *Publ. Dom. Ap. Obs.*, **6**, 134, 1933.

A STELLAR MODEL WITH A MIXED OPACITY LAW

C. M. Bondi

(Received 1949 May 26)

Summary

Stellar models are integrated in which the opacity is due both to photo-electric effect and to electron scattering. The integrations are performed by applying a method outlined in an earlier paper.*

---

1. In an earlier paper a method has been described, by the use of which the equations of stellar structure may be integrated by means of variables which are homology-invariant (h-variables). This method singles out the hydrostatic equation as the most important condition and one which is expressible immediately in terms of h-variables, while it places the energy transport equation in a secondary position (as it is physically). In the most important simple cases, the energy transport equation is expressible in terms of h-variables and the methods of Paper I are applicable. There is, however, a slightly more complicated case of considerable interest, mentioned briefly in Paper I, which will be discussed here. The energy transport equation is no longer expressible in terms of h-variables, but involves one variable which, while not homology-invariant, is very slowly varying. This case can be brought within the scope of the earlier method by slight modifications and approximations.

The model used in this paper is the generalized Cowling model, which consists of a core in convective equilibrium with an atmosphere in radiative equilibrium, radiation pressure being neglected throughout and the molecular weight being assumed to be constant. Energy generation is assumed to take place in the core only. The particular case mentioned above and to be discussed in this paper is that in which, in the region in radiative equilibrium, both the free-free transitions of electron scattering and the photoelectric effect make contributions to the opacity. The guillotine factor in the photoelectric effect is neglected. This is a case of considerable physical interest, and does not appear to have been discussed elsewhere. It should be noticed that the relative contribution of the two sources to the opacity varies from point to point, and the contribution from electron scattering always increases relatively from the surface towards the centre. It can be shown from the equations that at low temperatures, that is, very near the star's surface, the contribution of photoelectric effect to the opacity (Kramers' opacity) always completely outweighs that due to electron scattering. Thus a radiative atmosphere with opacity due purely to electron scattering does not in fact exist, although such an atmosphere may exist with opacity due purely to photoelectric effect. It is thus of considerable physical interest to investigate a whole range of possible configurations from the case of pure Kramers' opacity to those in which, at some point interior to the surface but close to it, electron scattering opacity is predominant. Since, at the surface, Kramers' opacity must always predominate, it is models of this last type which are in fact of interest when

* C. M. Bondi and H. Bondi, *M.N.*, **109**, 62, 1949. This will be referred to as Paper I.

investigating stars in which, through a large proportion of the volume, electron scattering opacity predominates.

2. The contribution of photoelectric effect to the opacity is given by the expression $\alpha(1-X^2)\rho/T^{3\cdot5}$, where α is a constant, X is the proportion of hydrogen by mass, ρ is the density and T the temperature. The corresponding contribution of electron scattering is $\kappa(1+X)$, where κ is a constant.

The energy transport equation in a radiative region can be expressed conveniently in the form

$$\frac{dT}{dr} = \text{opacity} \times \frac{3L_s}{16\pi ac}\frac{\rho}{r^2T^3},$$

where L_s is the total luminosity * and the usual notation is used. It is convenient to introduce constants B and C, defined by

$$B = \frac{3\alpha(1-X^2)}{16\pi ac}L_s$$

and

$$C = \frac{3\kappa(1+X)}{16\pi ac}L_s.$$

B and C are constant in any one configuration but vary between configurations.

It has been pointed out above that the relative contribution of the two sources of opacity varies throughout the radiative region, so that no constant can express their ratio at all points of a configuration. However, it is clear that some parameter involving B and C can be defined so that all possible configurations form an ordered sequence, a member of the sequence being defined by a particular value of the parameter. A convenient form for this parameter was found to be

$$E = \frac{\mathscr{R}r_s^3B^6}{4\pi GC^{11}\mu},$$

where $\mathscr{R}$ is the gas constant, G the constant of gravitation, μ the molecular weight and r_s the radius of the star. The particular significance of E will be seen later.

It should be observed that when E and the chemical composition have been specified, the integrations lead to only a single infinity of possible stars, that is, given the mass, the radius and luminosity are determined. Thus, the assumption of a particular mode of generation of energy, such as Bethe's law, defines a single star of a particular mass, luminosity and radius.

3. We now consider the form of the equations. The hydrostatic equation is, in the usual notation,

$$\frac{dP}{dr} = -\frac{G\rho M}{r^2}, \tag{1}$$

while conservation of mass is given by

$$\frac{dM}{dr} = 4\pi r^2\rho. \tag{2}$$

The energy transport equation is

$$\frac{dT}{dr} = -\frac{B\rho^2}{r^2T^{6\cdot5}} - \frac{C\rho}{r^2T^3}. \tag{3}$$

Since radiation pressure is neglected,

$$P = \frac{\mathscr{R}\rho T}{\mu}. \tag{4}$$

* The suffix s refers to surface values.

The variables S, Q, N introduced in Paper I are defined by

$$S=\frac{4\pi Pr^4}{GM^2}, \qquad Q=\frac{Pr}{GM\rho}, \qquad 1-N=\frac{d(\log\rho)}{d(\log P)}.$$

As shown there, the hydrostatic and mass conservation conditions together yield the equation:

$$\frac{S}{Q}\frac{dQ}{dS}=\frac{S+N-Q}{1+2S-4Q}. \tag{5}$$

Since

$$\frac{r}{T}\frac{dT}{dr}=N\frac{r}{P}\frac{dP}{dr},$$

it follows that

$$\frac{N}{Q}=\frac{C\rho}{rT^4}+\frac{B\rho^2}{rT^{7.5}}$$

or

$$N=\frac{C}{G}\left(\frac{\mathscr{R}}{\mu}\right)^4\frac{\rho^4}{MP^3}(1+\sigma), \tag{6}$$

where

$$\sigma=\frac{B}{C}\left(\frac{\mathscr{R}}{\mu}\right)^{3.5}\frac{\rho^{4.5}}{P^{3.5}}. \tag{7}$$

σ is therefore the ratio of the contributions of Kramers' and electron opacities at any point. It is not an h-variable.

The energy transport equation now becomes

$$\frac{S}{N}\frac{dN}{dS}=\frac{S+(1-4N)-\dfrac{\sigma}{1+\sigma}(4{\cdot}5N-1)}{1+2S-4Q} \tag{8}$$

or, putting $\sigma/(1+\sigma)=\nu$,

$$\frac{S}{N}\frac{dN}{dS}=\frac{S+(1-4N)-\nu(4{\cdot}5N-1)}{1+2S-4Q}. \tag{8'}$$

From the definition of E and of σ, it follows that

$$\frac{\nu^6}{(1-\nu)^{11}}=E\frac{SN^5}{Q^7}\left(\frac{r}{r_s}\right)^3. \tag{9}$$

This brings out the significance of E and also emphasizes the fact that ν is not an h-variable, for on the right-hand side of equation (9) S, N, Q, but not r/r_s, are h-variables.

From its definition, it is necessary that $0\leqslant\nu\leqslant 1$. $\nu=0$ corresponds to opacity due entirely to electron scattering, and $\nu=1$ corresponds to opacity due entirely to photoelectric effect. On the other hand, $\nu^6/(1-\nu)^{11}$ can vary between 0 and infinity, and hence the right-hand side of equation (9), which can be treated as a function of S, can range from 0 to infinity. Hence ν is a slowly varying function of S. Or, to express it otherwise, if the right-hand side of (9) is known approximately, ν can be determined accurately. For example, if $E[SN^5(r/r_s)^3/Q^7]$ is equal to $1{\cdot}09\times10^3$, then ν is equal to 0·60, whereas if $E[SN^5(r/r_s)^3/Q^7]$ is equal to $1{\cdot}66\times10^3$, then ν has only changed to 0·61.

We now discuss briefly the point which has been mentioned above, that at the surface the opacity is due entirely to photoelectric effect.

In the neighbourhood of the surface, that is, as T tends to zero, it is necessary that ρ should vary like $T^{3.25}$. Hence, the second term on the right-hand side of

equation (3) which arises from electron scattering, is proportional to $T^{0\cdot25}$, while the first term, which arises from photoelectric effect, is constant. Therefore, as T tends to zero, the photoelectric effect term becomes relatively large. This is equivalent to the statement that the surface condition on σ is that σ tends to infinity as r tends to r_s, or that ν tends to 1 as r tends to r_s.

4. The problem therefore becomes that of integrating the equations (5) and (8′) with S as independent variable, where ν is given by (9). The problem could be solved in exactly the same manner as outlined for various other problems in Paper I but for the fact that r/r_s enters into (9).

It might be suggested that a differential equation might be used to determine ν, since in fact ν satisfies the equation

$$\frac{S}{\nu(1-\nu)}\frac{d\nu}{dS} = -\frac{4\cdot5N-1}{1+2S-4Q}, \tag{10}$$

which, apart from ν, contains only h-variables. However, the simplicity of this method is somewhat illusory, since a solution for ν must involve a constant of integration which is not homology-invariant.

In fact it was found that since complete integrations were available for the cases of opacity due purely to photoelectric effect and due purely to electron scattering, r/r_s could be reasonably well estimated for intermediate cases. Since, as has been pointed out, $\nu^6/(1-\nu)^{11}$ is a very quickly varying function of ν, this was sufficient to give an accurate estimate of ν.

Since the neighbourhood of the star's surface is of greatest interest in the integration, the changeover from the mixed opacity law to the Kramers' opacity law occurring there, the integration was performed in two parts; for convenience the dividing line being at $S=0\cdot1$. First, a value of E was chosen. To each value of E there corresponds a single infinity of solutions running to the origin, and any such solution can be characterized by the value of Q at some specified value of S, say $S=0\cdot1$. For any such solution, N is then completely determined as a function of S. The required solution is that which gives the correct conditions of fit on to the convective core, the solution for which is known, being an incomplete Emden polytrope for $n=\frac{3}{2}$. Some value of Q at $S=0\cdot1$ was therefore assumed, and from this assumption the value of N at $S=0\cdot1$ for this solution was evaluated. In fact, an estimate of N at $S=0\cdot1$ must be made, but this can be made very accurately, since a good approximation can be made by interpolation between the two extreme cases (integrations in these two cases already existing). This approximation can then be tested by the smoothness of the resulting integration of N as a function of S, which is highly stable. The value of ν at $S=0\cdot1$ can be obtained from the values of Q and N there and from an estimate of r/r_s.

The integration of the equations for increasing S (i.e. inwards) then proceeded exactly as in Paper I, using the graphs described there, except that ν had to be evaluated at each stage from an assumed value of r/r_s. If no fit on to the core was obtained, another value of Q was taken and the integration repeated. The procedure was continued until the correct value of Q at $S=0\cdot1$ was found. In practice, the correct value of Q at $S=0\cdot1$ was not difficult to find, as it can be estimated by interpolation between the two extreme cases. It was generally found that not more than two integrations had to be performed.

Having determined, for a given value of E, the value of Q at $S=0\cdot1$ which gives a point of fit on to the Emden core, we consider the integration outwards

from $S=0{\cdot}1$ to the surface. Integration in this direction is highly unstable for N and the most satisfactory procedure is as follows. At $S=0{\cdot}0$ (the surface) $N=\frac{4}{17}$, while at $S=0{\cdot}1$ a good estimate of it has been obtained. Also at the surface dN/dS is infinite. The run of N must lie between the two extreme cases and therefore some idea of the correct run of N can be obtained. A point which assists this construction is that the value of E gives some indication of whether the correct solution is close to the Kramers' opacity solution or the electron opacity solution. In the former case, the construction of the run of N is straightforward, since the surface boundary condition for N is the same, and the first approximation is frequently sufficient. In the latter case, more care is required and higher approximations may be necessary. That is, when a complete solution has been obtained from the assumed run of N, the resulting dN/dS must be compared with the assumed dN/dS. If there is any considerable discrepancy, a new run of N must be assumed and the integration repeated.

Once a run of N has been assumed for $0\leqslant S\leqslant 0{\cdot}1$ only equation (5) remains to be integrated. In this region the use of the graphs of Paper I is not appropriate, and the equation is best solved in the form

$$\frac{d\log Q}{d\log S}=\frac{S+N-Q}{1+2S-4Q}. \tag{5'}$$

As S decreases, $\log Q$ and hence Q can be obtained at each stage. r/r_s can then be obtained from the quadrature

$$\log r/r_s=-\int_0^Q \frac{dQ}{S+N-Q}. \tag{11}$$

Hence ν can be calculated and dN/dS. This is compared with the assumed dN/dS and if necessary the integration is repeated with the new run of N assumed. This second run of N has in most cases been found to be unnecessary, but has been required and been found sufficient in the cases nearest to pure electron scattering opacity.

5. In order to make full use of the completed integration, we must obtain the relations between the mass, luminosity and radius of the star, given its composition. These involve knowing the ways in which Q and ν vary with S at the surface. Since $N=\frac{4}{17}$ at $S=0$, it is necessary that in the neighbourhood of the surface,

$$Q=AS^{4/17},$$

where A is a constant which can be determined from the integration. Also, from (8') and (10),

$$N=\tfrac{4}{17}\left[1+\frac{\delta}{35}S^{1\,17}\right]$$

and

$$\nu=1-\delta S^{1/17}$$

in the neighbourhood of the surface, where δ is given by

$$\delta^{11}=(\tfrac{17}{4})^5\frac{A^7}{E}. \tag{12}$$

From (6) and the definitions of S, Q, N, it follows that

$$\frac{C}{\pi M^3}\left(\frac{\mathscr{R}}{\mu G}\right)^4=\frac{4N(1-\nu)Q^4}{S}, \tag{13}$$

while, multiplying (6) by (7), we have

$$\frac{Br^{0\cdot5}}{\pi^2M^{5\cdot5}}\left(\frac{\mathscr{R}}{\mu G}\right)^{7\cdot5}=\frac{16\nu Q^{8\cdot5}N}{S^2}. \tag{14}$$

Consider these equations at the surface.

Then

$$\frac{C}{\pi M_s^3}\left(\frac{\mathscr{R}}{\mu G}\right)^4=\tfrac{16}{17}\delta A^4=K_1 \quad \text{say,} \tag{15}$$

and

$$\frac{Br_s^{0\cdot5}}{\pi^2M_s^{5\cdot5}}\left(\frac{\mathscr{R}}{\mu G}\right)^{7\cdot5}=\tfrac{64}{17}A^{8\cdot5}=K_2 \quad \text{say.} \tag{16}$$

Recalling the definitions of B and C, these are the required relations between M_s, r_s and the luminosity L_s, which enters through B and C. It is therefore essential to evaluate A and hence δ.

One method of evaluating A is directly from the integration, that is by continuing the outward integration for decreasing S until $(\log Q-\frac{4}{17}\log S)$ becomes constant. This is extremely tedious and inconclusive and another method presents itself, which is much more satisfactory.

Consider again equation (15) and divide it by (13). Then

$$\left(\frac{M}{M_s}\right)^3=\frac{4\delta A^4}{17}\frac{S}{N(1-\nu)Q^4}$$

$$=(\tfrac{4}{17})^{6/11}\frac{S}{N(1-\nu)Q^4E^{1/11}}A^{51\,11} \quad \text{from (12).} \tag{17}$$

But M/M_s can also be obtained by a quadrature, which can very easily be performed. For

$$\log\frac{M}{M_s}=-\int_0^S\frac{dS}{1+2S-4Q}. \tag{18}$$

We can therefore evaluate M/M_s at some arbitrary point S_1 from (18) and so find A from (17).

P/P_c*, ρ/ρ_c and T/T_c were then calculated as in Paper I. The results of these integrations are given in Table I.

TABLE Ia

E infinite, i. e. *opacity entirely due to photoelectric effect.* $\nu=1$ *throughout*

At the surface $Q=0\cdot2351S^{4/17}$

S	Q	N	r/r_s	M/M_s	P/P_c	ρ/ρ_c	T/T_c
0	0	0·2353	1	1	0	0	0
0·1	0·144	0·2443	0·459	0·854	0·0164	0·0553	0·297
0·2	0·184	0·2533	0·387	0·735	0·0481	0·124	0·387
0·3	0·220	0·2621	0·344	0·637	0·0869	0·192	0·453
0·5	0·286	0·2791	0·293	0·498	0·168	0·312	0·538
0·7	0·352	0·2954	0·260	0·400	0·245	0·408	0·600
1·0	0·451	0·3190	0·228	0·302	0·337	0·509	0·662
1·4	0·581	0·3488	0·201	0·225	0·434	0·601	0·722
2·0	0·780	0·3917	0·174	0·157	0·537	0·688	0·781
2·12	0·820	0·4000	0·170	0·147	0·548	0·697	0·786

The convective core fits on at $S=2\cdot12$, the radius of the core in Emden units being 1·19.

* The suffix c refers to central values.

TABLE Ib

log E=9·98. *At the surface,* $1-\nu=0{\cdot}0943S^{1/17}$, $Q=0{\cdot}2335S^{4/17}$

S	Q	N	ν	r/r_s	M/M_s	P/P_c	ρ/ρ_c	T/T_c
0	0	0·2353	1	1	1	0	0	0
0·1	0·143	0·2452	0·902	0·459	0·869	0·0169	0·0570	0·296
0·2	0·183	0·2548	0·890	0·388	0·750	0·0492	0·128	0·384
0·3	0·218	0·2641	0·881	0·345	0·653	0·0898	0·199	0·451
0·5	0·284	0·2824	0·873	0·296	0·510	0·169	0·316	0·535
0·7	0·350	0·3000	0·850	0·263	0·409	0·243	0·410	0·593
1·0	0·448	0·3256	0·833	0·231	0·311	0·338	0·513	0·659
1·4	0·580	0·3586	0·813	0·204	0·230	0·425	0·596	0·713
1·92	0·753	0·4000	0·792	0·1795	0·168	0·519	0·675	0·769

The convective core fits on at $S=1{\cdot}92$, the radius of the core in Emden units being 1·245.

TABLE Ic

log E=0·5055. *At the surface,* $1-\nu=0{\cdot}645S^{1/17}$, $Q=0{\cdot}2113S^{4/17}$

S	Q	N	ν	r/r_s	M/M_s	P/P_c	ρ/ρ_c	T/T_c
0	0	0·2353	1	1	1	0	0	0
0·1	0·140	0·2500	0·500	0·476	0·857	0·0168	0·0590	0·285
0·2	0·180	0·2629	0·460	0·404	0·741	0·0483	0·130	0·372
0·3	0·215	0·2750	0·435	0·362	0·649	0·0862	0·198	0·435
0·5	0·282	0·2988	0·387	0·309	0·507	0·165	0·317	0·521
0·7	0·346	0·3221	0·354	0·277	0·414	0·238	0·409	0·582
1·0	0·444	0·3565	0·314	0·243	0·317	0·338	0·517	0·654
1·38	0·571	0·4000	0·277	0·215	0·237	0·426	0·599	0·711

The convective core fits on at $S=1{\cdot}38$, the radius of the core in Emden units being 1·415.

TABLE Id

log E=$\overline{14}$·6037. *At the surface,* $1-\nu=8{\cdot}56S^{1/17}$, $Q=0{\cdot}1266S^{4/17}$

S	Q	N	ν	r/r_s	M/M_s	P/P_c	ρ/ρ_c	T/T_c
0	0	0·2353	1	1	1	0	0	0
0·04	0·0990	0·2555	0·0125	0·618	0·946	0·0052	0·0268	0·194
0·1	0·132	0·2650	0·0100	0·530	0·866	0·0201	0·0730	0·275
0·2	0·170	0·2799	0·0082	0·460	0·764	0·0560	0·154	0·357
0·3	0·204	0·2957	0·0071	0·417	0·674	0·0954	0·227	0·420
0·5	0·273	0·3249	0·0055	0·360	0·526	0·177	0·343	0·516
0·7	0·340	0·3538	0·0046	0·324	0·426	0·244	0·427	0·571
1·0	0·443	0·3960	0·0037	0·287	0·325	0·333	0·517	0·644
1·03	0·453	0·4000	0·0037	0·284	0·318	0·343	0·526	0·652

The convective core fits on at $S=1{\cdot}03$, the radius of the core in Emden units being 1·58.

TABLE Ie

$E=0$, i.e. mathematically limiting case of opacity due entirely to electron scattering $\nu=0$ throughout. At the surface, $Q=0{\cdot}218S^{\frac{1}{2}}$

S	Q	N	r/r_s	M/M_s	P/P_c	ρ/ρ_c	T/T_c
0	0	0·2500	1	1	0	0	0
0·04	0·101	0·2559	0·615	0·947	0·005	0·027	0·196
0·1	0·132	0·2661	0·532	0·888	0·020	0·073	0·274
0·2	0·170	0·2808	0·462	0·784	0·056	0·154	0·357
0·3	0·204	0·2958	0·419	0·691	0·095	0·226	0·420
0·5	0·273	0·3254	0·361	0·539	0·177	0·343	0·516
0·7	0·340	0·3541	0·325	0·437	0·245	0·429	0·571
1·0	0·443	0·3960	0·288	0·333	0·333	0·518	0·643
1·03	0·453	0·4000	0·285	0·326	0·343	0·527	0·652

The convective core fits on at $S=1{\cdot}03$, the radius of the core in Emden units being 1·58.

In order to investigate the changes in the mass-luminosity-radius relation, it is of interest to tabulate K_1 and K_2 against E. The expressions K_1 and K_2 are non-dimensional, and it is also of interest to determine corresponding dimensional quantites defined by

$$\frac{\mu^4}{(1+X)}\frac{M_s^3}{L_s} \quad \text{and} \quad \frac{\mu^{7\cdot5}}{(1-X^2)}\frac{{}^5_sM^{\cdot5}}{r_s^{0\cdot5}L_s}.$$

TABLE II

E	$\bar{\rho}/\rho_c$	$\frac{\mathcal{R}r_sT_c}{\mu GM_s}$	$\frac{\mu^4M_s^3}{L_s(1+X)}$	K_1	$\frac{\mu^{7\cdot5}M_s^{5\cdot5}}{r_s^{0\cdot5}(1-X^2)L_s}$	K_2	$K_1+10K_2^{6/11}$
0	0·049	0·80	[64·2168]	[$\bar{3}$·3538]	∞	0	[$\bar{3}$·3538]
[$\bar{1}\bar{4}$·6037]	0·0496	0·782	[64·2543]	[$\bar{3}$·3163]	[147·1514]	[$\bar{8}$·9462]	[$\bar{3}$·5431]
[0·5055]	0·0310	0·895	[64·4879]	[$\bar{3}$·0827]	[145·2603]	[$\bar{6}$·8372]	[$\bar{2}$·2172]
[9·9800]	0·0273	0·913	[65·1492]	[$\bar{4}$·4213]	[144·8913]	[$\bar{5}$·2062]	[$\bar{2}$·3899]
∞	0·0270	0·901	∞	0	[144·8661]	[$\bar{5}$·2314]	[$\bar{2}$·3989]

These are given in Table II. It can be seen that these can be compared respectively with the corresponding quantities for opacity due entirely to electron scattering, and opacity due entirely to photoelectric effect.

The non-dimensional quantities

$$\frac{\rho_{\text{mean}}}{\rho_c} \quad \text{and} \quad \frac{\mathcal{R}T_cr_s}{\mu GM_s}$$

are also given and can be compared with the extreme cases. For the sake of comparison, the extreme cases are also given in the tables, being the solutions given in Paper I.

Since K_1 is zero for infinite E and K_2 is zero for $E=0$, while $4EK_1^{11}=K_2^6$, for purposes of interpolation for various values of E it is convenient to tabulate $K_1+10K_2^{6/11}$ against E. Various combinations of this type have been examined but this was found to be the most satisfactory. Thus, for any E, it is possible by interpolation to evaluate K_1 and K_2.

It will be seen from these tables that the internal arrangements of stars change over smoothly from one extreme case, through the intermediate cases, to the other extreme case. The ratio of core radius to the radius of the star and the central condensation seem to vary monotonically from one extreme to the other, though there appears to be a slight waviness in the expression for T_c in terms of M_s and r_s. Columns 4 and 6 of Table II show how the addition of another source of opacity diminishes the luminosity. The infinity symbol occurring in these columns arises since, with fixed coefficients of opacity, pure electron opacity corresponds formally to a star of infinite mass, and pure Kramers' opacity to a star of zero mass. Now since for electron opacity $L_s \sim M_s^3$, $M_s^{5\cdot5} L_s^{-1}$ tends to infinity as M_s tends to infinity. The other infinity arises in a similar way and also has only formal significance.

Table III gives the detailed solutions in h-variables for the lower values of E for $S \leqslant 0\cdot1$. These are of interest, since the changeover from Kramers' opacity to mixed opacity takes place there for small E.

TABLE IIIA

Solution for small values of S for log E = 0·5055

S	Q	N	ν
0·0001	0·02560	0·23555	0·676
0·001	0·04385	0·23611	0·634
0·01	0·07545	0·2380	0·575
0·1	0·140	0·2500	0·500

TABLE IIIB

Solution for small values of S for log E = $\overline{14}$·6037

S	Q	N	ν
0·0001	0·02174	0·2370*	0·0298
0·001	0·03819	0·2478	0·0230
0·01	0·0682	0·2508	0·0163
0·1	0·132	0·2650	0·0100

* Owing to the instability of N when integrated outwards, no great accuracy is claimed for this figure, but Q and ν are accurately determined.

Newnham College,
Cambridge:
1949 *May* 26.

A COMPARISON OF THE SYSTEMS OF THE GENERAL CATALOGUE AND THE DRITTER FUNDAMENTALKATALOG

A. Kopff *

(Received 1949 July 22)

Summary

A discussion is made of the systematic errors of various fundamental catalogues, particularly the GC and FK3. It is concluded that the FK3 is on the whole at least as accurate as, if not superior to, the GC. For the present it appears advisable to use these two systems alongside each other, and not to try to combine them.

In 1943 Professor J. H. Oort wrote a paper (**1**) in which he made a number of comments on the system of the *Dritter Fundamentalkatalog* (FK3), and in particular compared the proper motions of the FK3 with those of the *General Catalogue* (GC). That paper gives rise to the remarks which follow; they would have appeared earlier, except that I did not learn of the investigation until the summer of 1948.

Professor Oort considers that the proper motions in declination are systematically of greater weight in the FK3 than in the GC, in respect of their dependence on both declination $(\mu_\delta)_\delta$, and right ascension $(\mu_\delta)_\alpha$; for the proper motions in right ascension he also prefers the FK3 so far as the $(\mu_\alpha)_\alpha$ are concerned, but he considers that the GC is the more reliable for the system of $(\mu_\alpha)_\delta$. The principal reason which he gives is that in deriving the dependence of the proper motions in right ascension on declination $(\mu_\alpha)_\delta$, part of the material consisted merely of short series of meridian circle observations.

Before considering this last point in detail, I propose to survey briefly the investigations undertaken for the systematic correction of Auwers' *New Fundamental Catalogue* (NFK). The papers are scattered widely in the *Astronomische Nachrichten*; but the essential points have been summarized and published, together with a bibliography (**2**). The manuscripts and calculations were destroyed during the war, so that it is no longer possible to amplify these accounts. The particular points of importance are as follows:—

(1) So far as the corrections to the separate stars in their original systems are concerned, the GC and FK3 are based in general on the same observational material. Each is also based on corrections to an existing catalogue: the GC on the correction and extension of the PGC of Lewis Boss (using the system of those stars regarded as principal standards in the strict sense in *A.J.* (**23**)); the FK3 on the correction and ultimate enlargement of Auwers' NFK (**3**). There is, however, one important difference in the manner of the corrections, which must be kept in mind in judging the result.

Analysis has shown that the systematic departures of the NFK places from those given by modern observations are primarily due to errors in the proper motions; the NFK places were essentially free from systematic errors at the time of its mean epoch. Errors in the NFK proper motions are due almost exclusively to the systematic errors of the earliest observational catalogues used by Auwers. This is not a hypothesis; it is manifest on comparison of these catalogues with the NFK, as is shown by the various studies in the *Astronomische Nachrichten* which will be considered below. Hence, when deriving the systematic corrections

* Associate of the Royal Astronomical Society.

to the NFK proper motions, use was made of only the more recent catalogues (starting with Pulkovo 1845), which could be regarded as essentially free from systematic instrumental errors. This rejection of the earliest catalogues represents a major advantage for the FK3, as compared with the system of the GC, for its proper motions are not influenced by the systematic errors of the oldest catalogues.

(2) Modern catalogues show that in right ascension and declination both the PGC and NFK have errors depending on right ascension. For the NFK the values of $(\Delta\alpha)_\alpha$ are tabulated for the northern hemisphere in (**4**), and for the southern in (**5**); the values of $(\Delta\delta)_\alpha$ are tabulated in (**6**). The sign of these corrections is opposite at the present time to what it was in about 1850. A fresh comparison was not carried out for catalogues observed before 1850, but such a comparison could be made by utilizing the work of Auwers (**7**).

The tables show consistently that the NFK requires a correction depending on right ascension, to both μ_α and μ_δ. The fact that the catalogues from about 1850 onwards show a reversal of sign in $(\Delta\alpha)_\alpha$ and $(\Delta\delta)_\alpha$ is conclusive proof that the erroneous proper motions of the NFK arise through inclusion of catalogues previous to 1850. The earliest catalogues used by Auwers for the NFK (**7**) contain systematic errors due mainly to errors of clock rate, level and azimuth, and to erroneous refraction corrections; these enter into the proper motions and thus led to the observed deviations of the NFK after 1845. Attention should also be drawn to the reversal in the $(\Delta\delta)_\alpha$ for northern catalogues at the zenith (**8**). In fact, better proper motions would have been obtained for the NFK if the oldest catalogues had been altogether ignored; admittedly, however, the total number of available catalogues was so small that this could hardly have been done. If we take the modern corrections $(\Delta\mu_\alpha)_\alpha$ and $(\Delta\mu_\delta)_\alpha$ to the NFK, and use them to calculate the date for which the present-day $(\Delta\alpha)_\alpha$ and $(\Delta\delta)_\alpha$ of the NFK reduce on the whole to zero, we get in fact approximately the date of mean epoch of the NFK places. (For R.A. compare (**9**) and the curves given therein; for declination (**10**).)

One is forced to conclude from this that the NFK at the date of its mean epoch was free from major systematic errors of the type $(\Delta\alpha)_\alpha$ and $(\Delta\delta)_\alpha$, but that errors of its proper motions were caused by the inclusion of catalogues earlier than 1845. Thus one can obtain good values for the corrections to the NFK proper motions simply by taking the $(\Delta\alpha)_\alpha$ and $(\Delta\delta)_\alpha$ determined for it at the present day, and dividing them by the time since its mean epoch.

This is immediately evident from the equations used to obtain the corrections to the places and proper motions. With a series of equations of condition of the form $x+yt=l$ (weight p), and with $[pt]=0$,

$$x=\frac{[pl]}{[p]}, \qquad y=\frac{[ptl]}{[pt^2]}.$$

The mean correction in position at the epoch given by the condition $[pt]=0$ is independent of t. Since the weights of the earliest catalogues were assessed very low by Auwers, these catalogues have only a minor influence on the places of the NFK at the time of its mean epoch; but their errors count heavily in the proper motions y.

The corrections $(\Delta\mu_\alpha)_\alpha$ and $(\Delta\mu_\delta)_\alpha$ of the NFK have received particular attention on account of their importance for studies of the systematic motions within the galactic system (**11**).

In forming the GC similar corrections were not undertaken (**12**); the errors arising from the earliest catalogues are thus still present in the GC proper motions.

The considerations which have been set forth *in extenso* are equally applicable to a discussion of $(\Delta\alpha)_\delta$ and $(\Delta\delta)_\delta$. The argument will not be repeated in detail; reference to the relevant investigations in the *Astronomische Nachrichten* will suffice, and in particular attention is drawn to Section (4) below.

(3) The procedure used in obtaining the proper motion corrections $(\Delta\mu_\delta)_\delta$ is different for the two fundamental systems. The GC uses a sort of interpolation method. Declinations observed in both the northern and the southern hemispheres are brought into agreement over their common range by an adjustment of the constant of refraction for each catalogue; the declination systems of the individual observational catalogues, corrected in this way, are then used for the corrections $(\Delta\mu_\delta)_\delta$. Attempts of this kind, made during the correction of the NFK, led in some cases to such improbable corrections for the constant of refraction (**13**) that this method was abandoned. The discrepancies between north and south catalogues for the same region are created more by instrumental errors, especially flexure, than by uncertainty in the refraction. The procedure of the GC is thus without a physical basis: it is a pure interpolation method, but one which may in some circumstances yield quite serviceable results.

In contrast to this, the corrections $(\Delta\mu_\delta)_\delta$ of the NFK were based on observations of the Sun, and to some extent of Mercury and Venus also (**14**). The adequacy or inadequacy of these corrections will require some considerable time to establish. Oort and others have concluded that in this respect the FK3 is either equal or superior to the GC.

Again, in the determination of the $(\Delta\mu_\delta)_\delta$ of the NFK it was found that catalogues between 1850 and the present day show a regular variation in the differences $\Delta\delta_\delta$ with respect to the NFK. In each individual range of declination there is a run in $\Delta\delta_\delta$ with time, which gives, if one excludes the earliest catalogues (i. e. those before Pulkovo of epoch 1846), a very satisfactory correction $(\Delta\mu_\delta)_\delta$ (**15**). Further, calculation of the date for which $(\Delta\delta)_\delta$ for epoch 1925 of the NFK vanish as a result of the corrections $(\Delta\mu_\delta)_\delta$, gives the year 1864 for the northern hemisphere and 1866 for the southern. The NFK, therefore, at the date of its mean epoch was on the whole free from systematic errors in declination, of the form $(\Delta\delta)_\delta$; here again one could infer the corrections to the NFK proper motions simply from the $(\Delta\delta)_\delta$ of the present day, thus essentially eliminating the earliest catalogues.

(4) We have now to deal with the problem of the corrections $(\Delta\mu_\alpha)_\delta$. For the GC the errors $(\Delta\mu_\alpha)_\delta$ have been derived from all observational catalogues whose systems could be regarded as independent. The method used for the correction of the NFK was a different one.

In this case the corrections for the northern hemisphere down to the equator must be considered separately from those of the southern hemisphere. First of all (**16**), we must take into consideration the condition that for the determination of $(\Delta\mu_\alpha)_\delta$, only those observational catalogues may be used for which the instrumental system α_δ can be regarded as constant over a series of epochs. For this purpose the paramount need is for a series of catalogues depending on observations with the same unchanged instrument, and with careful determination of the azimuth. Material of this kind is very scanty. It has been found that the only series of observations which on the whole fulfils these conditions is that obtained in Pulkovo between 1845 and 1915, for declinations 0° to +50°; but the extended

series obtained at Greenwich reproduce satisfactorily at least the relative run of the $(\Delta\mu_\alpha)_\delta$. That the $(\Delta\mu_\alpha)_\delta$ derived from the Pulkovo series alone gives satisfactory systematic corrections to the system of the NFK can be seen from the fact that the $(\Delta\alpha)_\delta$ of the NFK can also be represented by these corrections (**17**). It would thus have been sufficient for the derivation of the $(\Delta\mu_\alpha)_\delta$ in the region 0° to +50° to divide the $(\Delta\alpha)_\delta$ by the epoch difference 1864–1910. In the declination region north of +50° the correction $(\Delta\mu_\alpha)_\delta$ has been derived simply from the $(\Delta\alpha)_\delta$; in doing this for the region +50° to +70° the run of the $(\Delta\mu_\alpha)_\delta$ was determined from the Greenwich observations alone (**18**). The fact that we have restricted ourselves to the Pulkovo series only for the determination of the $(\Delta\mu_\alpha)_\delta$ in the range of declination from 0° to +50° is thus justified by the corresponding run of the $(\Delta\alpha)_\delta$. A sufficient correction $(\Delta\mu_\alpha)_\delta$ for the whole northern hemisphere would have been obtained, as we may say once more, by dividing the corrections $(\Delta\alpha)_\delta$ to the NFK by the difference of epoch between the NFK and FK3. Both the NFK and the FK3 are at the times of their respective epochs essentially free of systematic errors of the form $(\Delta\alpha)_\delta$. The systematic errors of the NFK at a later time are exclusively due to the systematically erroneous proper motions in the NFK.

In the southern sky the correction $(\Delta\mu_\alpha)_\delta$ was simply based on the $(\Delta\alpha)_\delta$ by multiplying by $\frac{100}{45}$ (**19**). It was therefore assumed that the corrections $(\Delta\alpha)_\delta$ of the NFK in the southern hemisphere have arisen from corresponding errors of the proper motions during these 45 years. The extensive investigations of A. Kahrstedt (**5**) served to justify this assumption. No long series of observations in the southern hemisphere corresponding to the Pulkovo series is available; even the longest Cape series extends only over 31 years, but it confirms essentially the assumption on which the $(\Delta\mu_\alpha)_\delta$ are based.

Calculating once more the time at which the $(\Delta\alpha)_\delta$ vanish, if their size is determined by the $(\Delta\mu_\alpha)_\delta$, we obtain 1860 for the northern sky, using Pulkovo (**20**), and about 1865 for the southern sky (**21**). We are therefore justified here in coming to the same conclusions as in the case of the other systematic corrections to the proper motions.

The following table gives the run of the differences of the $(\mu_\alpha)_\delta$ for the various catalogues. The arrangement is the same as that used by Oort. The figures are given in seconds of time and refer to the comparison FK3 – GC (**22**), FK3 – PGC and FK3 – NFK (both from a paper by H. Nowacki (**23**), using slightly smoothed values).

TABLE I

Systematic Differences in $(\mu_\alpha)_\delta$ per century

Dec.		FK3 – GC	FK3 – PGC	FK3 – NFK
		s	s	s
+90° to	+83°	−0·001 sec δ	−0·007 sec δ	−0·003 sec δ
+82	+68	−0·022 sec δ	−0·035 sec δ	−0·012 sec δ
+67	+53	−0·025 sec δ	−0·040 sec δ	−0·027 sec δ
+52	+38	−0·016	−0·033	−0·049
+37	+23	−0·015	−0·024	−0·039
+22	+ 8	−0·009	−0·008	−0·018
+ 7	− 7	−0·004	−0·004	−0·007
− 8	−22	+0·002	+0·005	+0·010
−23	−37	+0·015	+0·011	+0·061
−38	−52	+0·010	+0·025	+0·113
−53	−67	+0·009 sec δ	+0·044 sec δ	+0·050 sec δ
−68	−82	+0·014 sec δ	+0·073 sec δ	+0·049 sec δ
−83	−90	+0·009 sec δ	+0·045 sec δ	+0·030 sec δ

The last column gives the comparison FK3−NFK. These figures show that in general the agreement between the FK3 and the GC is closer than that between either of them and the PGC or NFK. In fact the alteration from the PGC to the GC is similar to that from NFK to FK3, though a little smaller in amount.

Now it can immediately be shown that the reason lies in the use of different raw material. When the corrections to the proper motions of the NFK were obtained, the oldest catalogues were omitted on account of their systematic uncertainties, but the GC on the other hand included these. For the four most northerly zones (+80°, +60°, +45°, +30°), the reductions mentioned in the *G.C.*, Vol. I, p. 19 have been repeated by Dr S. Böhme, neglecting the catalogues before 1845 in Table 7A, those before 1851 in Table 7B, as well as the catalogues in Table 7C and those values denoted as uncertain. The results are collected in the following Table II, giving for comparison the solution A + B from Table 8 of the GC.

TABLE II

Correction $(\Delta\mu_\alpha)_\delta$ *per century*

Dec.	GC	New
	s	s s
+80°	−0·098	−0·148 = −0·026 sec δ
+60	−0·038	−0·066 = −0·033 sec δ
+45	−0·018	−0·037
+30	+0·004	−0·025

Applying the new reductions of PGC to the values for FK3 − PGC in Table I, the amounts per century are reduced to those of Table III.

TABLE III

Dec.	FK3 − GC corr.
	s
+82° to +68°	−0·006 sec δ
+67 +53	−0·007 sec δ
+52 +38	+0·004
+37 +23	+0·001

The large quantities FK3 − GC have now almost completely vanished. The strong influence of the earliest catalogues on the proper motions is obvious on inspection of the tables of *G.C.*, Vol. I, pp. 17, 18. It was not necesssary for this purpose to carry through the calculations up to $\delta = -15°$. Here again a further approach of the GC to the FK3 can be seen as soon as the early catalogues are excluded. For the southern part of the southern sky so few catalogues are available that the determination of the systematic behaviour of the $(\Delta\mu_\alpha)_\delta$ is still very uncertain. The method which was used in deriving the FK3 (**5**) may also be preferable here, since the uncertain southern catalogues before about 1875 enter less into the proper motions of the FK3 than is the case with the GC.

The systematic corrections $(\Delta\mu_\alpha)_\delta$ of the NFK are based both in the northern part of the northern sky and in the southern sky on the whole material which was used when deriving $(\Delta\alpha)_\delta$ for the systematic correction of the NFK, excluding the oldest catalogues, which as the discussion has shown, are subject to large systematic errors.

(5) For the reasons given, the FK3 on the whole is at least as good as the GC, if not superior to it, in respect to the systematic behaviour of the proper motions. As to whether the one or the other system is the more valuable, only the distant future can decide; but it is probably of interest to consider the comparisons

W40 – FK3 and W40 – GC of the new catalogue of H. R. Morgan and F. P. Scott (**24**). The difference in time with respect to the mean epoch of the FK3 already amounts to 30 years. One may also compare the work of H. R. Morgan, "Corrections to the Fundamental Catalogs" (**25**). The difference in the means of $(\Delta\alpha)_\delta$ is caused by the difference of the equinoxes of the two fundamental catalogues. The $(\Delta\mu_\alpha)_\delta$ of the FK3 show in this work no dependence on δ and the corrections $(\Delta\mu_\alpha)_\delta$ of the GC are identical with the differences FK3 – GC in Table I, so that the differences in $\mu_\alpha \cos\delta$ in Oort's Table I (**1**), on the whole, are to be ascribed to the GC.

To mix in any way the two systems of the FK3 and GC can hardly be recommended, but rather one should use both systems alongside each other, following the example of A. N. Vyssotsky (**26**). After all, the system of the FK3 has formed the fundamental basis for all catalogue work, so that one is better advised to use it unaltered, and only resort to other systems for special researches.

Astronomisches Rechen-Institut,
Heidelberg:
1949 *July* 15.

References

(1) J. H. Oort, *B.A.N.*, **9,** No. 357, 1943.
(2) A. Kopff, *Veröff. Astr. Rechen-Inst., Berlin-Dahlem,* No. 54, 1937; see especially Appendix IV (A. Kahrstedt) and Appendix V (K. Heinemann).
(3) J. Peters, *Veröff. Astr. Rechen-Inst., Berlin-Dahlem,* No. 33, 1907.
(4) A. Kahrstedt, *A.N.*, **235,** 369, 1929.
(5) A. Kahrstedt, *A.N.*, **240,** 385, 1930.
(6) A. Kopff, H. Nowacki and K. Pilowski, *A.N.*, **245,** 49, 1932. Reprinted in *Mitt. Astr. Rechen-Inst., Berlin-Dahlem,* Bde. **2** and **3**.
(7) A. Auwers, *A.N.*, **164,** 225, 1904.
(8) A. Kopff, *A.N.*, **225,** 273, 1925. A. Kopff, H. Nowacki and K. Pilowski, *A.N.*, **245,** 52 and 60, 1932.
(9) A. Kahrstedt, *A.N.*, **235,** 393, 1929; *A.N.*, **240,** 396, 1930.
(10) A. Kopff, H. Nowacki and K. Pilowski, *A.N.*, **245,** 54 and 58, 1932.
(11) *E. g.* compare H. Nowacki, *A.N.*, **245,** 63, 1932; and J. H. Oort, *B.A.N.*, **9,** 425, 1943.
(12) B. Boss, *G.C.*, **1,** 11 and 24, 1937.
(13) K. Heinemann, *A.N.*, **241,** 151, 1930.
(14) K. Heinemann, *A.N.*, **249,** 289, 1933. K. Heinemann and H. Nowacki, *A.N.*, **252,** 353, 1934.
(15) K. Heinemann and H. Nowacki, see tables in *A.N.*, **252,** 379 and 386, 1934; and A. Auwers, *A.N.*, **164,** 225, 1904.
(16) A. Kahrstedt, *A.N.*, **235,** 369, 1929.
(17) *Ibid.*, p. 391.
(18) *Ibid.*, p. 399.
(19) A. Kopff, *Veröff. Astr. Rechen-Inst., Berlin-Dahlem,* No. 54, 108, 1937.
(20) A. Kahrstedt, *A.N.*, **235,** 392, 1929.
(21) A. Kahrstedt, *A.N.*, **240,** 396, 1930.
(22) A. Kopff, *A.N.*, **269,** 160, 1939.
(23) H. Nowacki, *A.N.*, **255,** 301, 1935.
(24) H. R. Morgan and F. P. Scott, *Publ. U.S. Naval Obs.*, **15,** Pt. 5, 1948.
(25) H. R. Morgan, *A.J.*, **54,** 145, 1949.
(26) A. N. Vyssotsky, *Publ. Leander McCormick Obs.*, **10,** No. 9, 1948.

ON THE EARLY APPEARANCE OF ATMOSPHERIC LINES IN THE ULTRA-VIOLET SPECTRUM OF ζ AURIGAE

Mario G. Fracastoro

(Communicated by the Director of the Cambridge Observatory)

(Received 1949 May 13)

Summary

A brief account is given of the early appearance of some lines in the ultra-violet spectrum of ζ Aurigae, at least 60 days before the beginning of the 1947–48 eclipse.

These lines belong mostly to *Fe* I, but also to *Ti* II, *Ca* II and probably *Sm* II and *Y* II.

Photometric measurements of the continuous ultra-violet spectrum show a general absorption produced by the gK5 star atmosphere, probably ten days before and eight days after the visual eclipse.

It may be recalled that the last eclipse of ζ Auri was predicted to begin on 1947 December 14·65. Our observations started about two months before this date, namely on 1947 October 15, with the aluminized 12-inch reflector of the Arcetri Observatory (Florence, Italy). The instrument was equipped with a 6°·1 prism, giving a dispersion of 126 A./mm. between H_5 and H_6 of the Balmer series. Fine-grain contrast plates were used, selected with the purpose of getting the highest resolving power, consistent with the size of the mirror and the magnitude of the star. The time of exposure ranged between 4 and 30 minutes. On each plate from 2 to 6 spectra were recorded with exposures varying usually in logarithmic progression. The spectra were widened to between 0·30 and 0·45 mm. The standardization of the plates has been done by means of the Pleiades; the spectra were mainly measured with the recording microphotometer of the Solar Physics Observatory at Cambridge.

These observations continued up to 1948 February 1, that is to about nine days after the end of the eclipse. From 1947 December 13 to 16, by kind permission of Father J. Stein, S.J., the 24-inch reflector of Castelgandolfo (Rome, Italy), with 2 prisms of 12° together has also been used.

The results obtained from the whole group of plates will be dealt with in a later communication. The aim of this paper is to announce the presence of a number of ultra-violet spectral lines (mainly beyond the last observed member, H_{12}, of the Balmer series) in all out-of-eclipse spectra of ζ Auri, dating back to 1947 October 15. The presence of such lines at this phase has never been published, as far as we know.

Some of these lines are readily seen on all the microphotometer records, provided that the exposure was sufficient for their particular spectral region. Other lines appear only in proximity to the eclipse. All are tabulated in Table I, together with their empirical intensities, deduced from the microphotometer tracings, and their identifications from Moore's Multiplet Tables (**1**). The numbers in brackets following the symbol of the element refer to the intensity and multiplet number, according to Moore's Tables.

Owing to the low resolving power of our instrument, other metallic lines belonging to the same multiplets are blended with some of the Balmer lines, namely:

H_5	λ3970·1 with	3968·5 *Ca* II.
H_7	3835·4 with	3834·2 *Fe* I (100) [20].
H_8	3797·9 with	3793·9 *Sm* II (500) [11].
H_{10}	3750·2 with	3749·5 *Fe* I (200) [21]; 3748·3 *Fe* I (60) [5]; 3745·9 *Fe* I (40) [5]; 3745·6 *Fe* I (100) [5].
H_{11}	3734·4 with	3737·1 *Fe* I (150) [5]; 3734·9 *Fe* I (300) [21]; 3731·3 *Sm* II (600) [11].
H_{12}	3721·9 with	3719·9 *Fe* I (250) [5].

It may appear puzzling that many observed lines belong to *Fe* I. These lines increase steadily in strength as the total phase is approached and are therefore fairly certainly low excitation lines. Furthermore, all the hydrogen lines between H_5 and H_{12} show an abnormally low central intensity. The only exceptions are H_6 and H_9, which are (as seen above) the only lines without any important blend. Table II gives the central intensities of these lines, referred to the continuous spectrum ($\lambda = 3950$) of the comparison star BD +41° 1044, $6^m{\cdot}41$, A5, which is very suitable for this purpose.

Although these lines are very broad, their central intensities might nevertheless be affected by the seeing. Consequently, the last column of Table II gives the depth of the K-line (p'_K) measured in the comparison star. This line is narrow and its depth is a very sensitive index of the general instrumental conditions of the plate, involving not only the seeing but also the optical performance of the photographic emulsion and the microphotometer.

The possibility is not excluded that the intensity of the hydrogen lines might increase according to the phase; but the rate of growth of the hydrogen density in the gK5-star atmosphere is certainly less important than the rate of growth of the metallic arc lines, as is shown by the central intensities of H_6 and H_9 as compared with the other members of the Balmer series. Typical, from this standpoint, is the behaviour of H_{10}, H_{11} and H_{12}, which remain very strong with their wavelength practically unchanged up to the maximum absorption phase, being gradually replaced by the above-mentioned *Fe* I lines.

It appeared very likely from Table I that the line intensities of plate 2 were abnormally high in comparison with the subsequent plates; the central intensities for plate 2, as seen in Table II, seem to confirm this result.

On the occasion of the previous eclipses of ζ Auri, many observers followed the behaviour of the K-line in order to get the first observational evidence of the presence of the atmosphere surrounding the gK5-component of the system. Beer (**2**) could observe this line up to 18 days after the eclipse in 1934 and 32 days before the eclipse in 1937, and deduced for the envelope a maximum radius r_a varying from 1·54 to 1·70, r_K being taken as unit.

The variable extent of the atmosphere was suggested also by Christie and Wilson (**3**). Further confirmation of this suggestion was given by Christie some years later (**4**). Other authors (**4, 5, 6**) suggested inhomogeneity of the atmosphere, consisting perhaps of temporary local condensations, analogous with solar prominences or coronal streamers. The peculiar behaviour of plate 2 (apart from the very early appearance of the lines) can be interpreted according to

TABLE I

Wave-lengths, empirical intensities and identifications of absorption lines in the ultra-violet spectrum of ζ Auri (from microphotometer recordings)

Plate	Epoch	Observed wave-lengths 3821*	3789	3760	3706·3	3688	3670	3663	3649	3633	3620	3606
	1947											
2	Oct. 15·9	...	...	...	3	0	1	...	2	4	...	2d
7	Oct. 20·8	...	...	...	2	0	...	...	1	2	...	1d
13	Nov. 4·9	1	...	...	0	...	...	...	1	1	...	...
14	Nov. 6·9	1	...	...	4	3	...	...	3	3	1	...
15	Nov. 7·9	...	...	...	3	1	...	...	2	2	...	0d
20	Nov. 20·9	1	...	...	3	2	...	...	1d	2	...	0
27	Dec. 5·0	0–1	...	0–1	2–3	2d	...	...	1d	1–2	2	2d
30	Dec. 7·8	0–1	0	1	3	3	...	...	2	3	...	...
64	Dec. 13·8	0	0–1	3	5	5d	...	...	5	5	2	4d
67	Dec. 14·77	0	0	4	8	6	...	2	7	6	3	8d
68	Dec. 14·84	1	...	4	8	6	2	2	9d	8	...	10d
69	Dec. 15·00	2d	1–2	5	8	6	1	0	6d	7	3	8d
	1948											
86	Jan. 30·9	1d	...	...	4	2d	0	...	3	4	...	1
96	Feb. 1·9	1d	...	...	2	1d	...	...	2	1	...	1d

3821 = 3820·4 *Fe* I (250) [20] ; 3815·8 *Fe* I (100) [45] ; 3825·9 *Fe* I (200) [20].
3789 = 3788·7 *Y* II (200) [7] ; 3790·1 *Fe* I (12) [22].
3760 = 3759·3 *Ti* II (200) [13] ; 3763·8 *Fe* I (100) [21] ; 3758·2 *Fe* I (150) [21].
3706·3 = 3706·0 *Ca* II (10) [3] ; 3706·2 *Ti* II (20) [73] ; 3705·6 *Fe* I (100) [5].
3688 = 3688·4 *Sm* II (100) [11] ; 3687·5 *Fe* I (40) [21].
3670 = 3670·8 *Sm* II (1000) [11] ; 3668·0 *Ce* II (400) [40].
3663 = 3662·2 *Ti* II (40) [75].
3649 = 3647·9 *Fe* I (100) [23].
3633 = 3633·1 *Y* II (200) [2] ; 3631·5 *Fe* I (125) [23] ; 3630·7 *Sc* II (50) [2].
3620 = 3618·8 *Fe* I (125) [23] ; 3619·4 *Ni* I (150) [35].
3606 = 3608·9 *Fe* I (100) [23] ; 3605·3 *Cr* I (750) [4].

* The 3821 band splits into three lines in the spectra with higher dispersion.

TABLE II

Central intensities of hydrogen and Ca II lines in log I (referred to the continuous spectrum of BD+41° 1044 at λ 3950 A.)

Plate	Epoch 1947	i_5	i_K	i_6	i_7	i_8	i_9	i_{10}	i_{11}	p'_K
2	Oct. 15·9	0·00	+0·19	−0·04	−0·19	−0·23	−0·19	−0·25	−0·28	14
7	Oct. 20·8	+0·03	+0·20	−0·01	−0·10	−0·14	−0·13	−0·18	−0·20	10
13	Nov. 4·9	−0·01	+0·19	−0·02	−0·14	−0·18	−0·15	−0·20	−0·24	15
14	Nov. 6·9	−0·02	+ 0·19	−0·04	−0·16	−0·21	−0·17	−0·25	−0·29	12
15	Nov. 7·9	−0·01	+0·18	−0·04	−0·14	−0·17	−0·16	−0·22	−0·24	13
16	Nov. 13·8	−0·03	+0·18	−0·04	−0·15	−0·17	−0·15	−0·23	−0·24	15
18	Nov. 17·8	0·00	+0·19	−0·01	−0·12	−0·13	−0·10	−0·19	−0·17	13
20	Nov. 20·9	−0·01	+0·19	−0·02	−0·14	−0·18	−0·17	−0·20	−0·25	14
26	Dec. 4·8	−0·11	+0·12	−0·07	−0·18	−0·18	−0·15	−0·22	−0·26	12
27	Dec. 5·0	−0·13	+0·10	−0·12	−0·21	−0·22	−0·21	−0·28	−0·29	12
30	Dec. 7·8	−0·13	+0·12	−0·10	−0·21	−0·24	−0·22	−0·32	−0·33	16
64	Dec. 13·8	−0·18	+0·01	−0·14	−0·28	−0·27	−0·27	−0·36	−0·39	13
	1948									
86	Jan. 30·9	−0·05	+0·18	−0·07	−0·17	−0·21	−0·18	−0·27	−0·27	15
96	Feb. 1·9	−0·02	+0·18	−0·06	−0·16	−0·19	−0·18	−0·25	−0·28	16

TABLE III *a*

Differences in log I between η Auri and ζ Auri at the same wave-lengths, deduced with correcting factors (n=number of spectra)

λ	Plate 2	n	7	n	13	n	14	n	20	n	27	n	30	n	64	n	86	n	96	n
A. 3672	1·12	3	...	...	1·12	1	1·10	2	1·20	2	1·12	1	...	...	1·13	2	...	...	...	...
B. 3634	1·13	3	1·09	2	1·11	1	1·08	2	1·18	2	1·21	1	...	...	1·17	2	...	...	1·12	2
C. 3600	1·18	3	1·08	2	1·13	1	1·09	2	1·20	2	1·24	1	...	...	1·20	2	...	...	1·16	2
D. 3560	1·15	3	1·12	2	1·16	1	1·04	2	1·18	2	1·21	1	1·30	1	1·26	2	1·14	1	1·16	2
E. 3530	1·14	3	1·10	2	1·16	1	1·05	3	1·17	2	1·19	1	1·25	1	1·27	2	1·15	1	1·12	2
F. 3500	1·12	3	1·08	2	1·18	1	1·05	3	1·16	2	1·19	1	1·26	1	1·31	2	1·13	1	1·10	2
G. 3470	1·17	3	1·04	2	...	...	1·02	2	1·20	2	1·19	1	1·26	1	1·34	2	1·15	1	1·06	2
H. 3440	1·16	2	1·08	2	...	...	...	...	1·13	1	1·16	1	...	...	1·28	1	1·15	1	1·07	1
I. 3410	1·20	2	...	...	...	...	...	...	...	...	...	...	...	...	1·33	1	1·11	1	...	...
Mean	1·15		1·08		1·14		1·06		1·18		1·19		1·27		1·25		1·14		1·11	

TABLE III *b*

Differences in log I between η Auri at λ 3410 A. and ζ Auri at various wave-lengths (n = number of spectra)

λ	Plate 2	n	7	n	13	n	14	n	15	n	20	n	27	n	30	n	64	n	86	n	96	n
A.	0·33	3	0·30	3	0·37	1	0·36	2	0·30	1	0·32	2	0·40	3	0·44	2	0·42	3	0·36	3	0·36	3
B.	0·41	3	0·36	3	0·40	1	0·43	2	0·40	1	0·39	2	0·45	3	0·48	2	0·51	3	0·45	3	0·44	3
C.	0·50	3	0·44	3	0·49	1	0·52	2	0·46	1	0·46	2	0·54	3	0·58	2	0·61	3	0·54	3	0·52	3
D.	0·57	3	0·52	3	0·60	1	0·60	2	0·55	1	0·54	2	0·61	3	0·70	2	0·70	3	0·64	3	0·60	3
E.	0·62	3	0·56	3	0·65	1	0·64	2	0·64	1	0·60	2	0·67	3	0·77	1	0·78	3	0·69	3	0·66	2
F.	0·69	3	0·64	3	0·74	1	0·73	2	0·73	1	0·68	2	0·74	3	0·86	1	0·90	3	0·78	3	0·74	2
G.	0·82	2	0·72	3	...	...	0·80	1	...	...	0·79	2	0·82	3	0·98	1	1·03	3	0·89	3	0·80	2
H.	0·89	2	0·90	2	...	...	...	...	...	...	0·91	1	0·90	2	1·10	1	1·13	2	0·98	2	0·92	1
I.	1·06	1	...	...	...	...	...	...	...	...	...	...	1·01	2	...	...	1·27	2	1·10	1	...	...
Mean A–E	0·49		0·44		0·50		0·51		0·47		0·46		0·53		0·59		0·60		0·54		0·52	
A–G	0·56		0·51		...		0·58		...		0·54		0·60		0·69		0·71		0·62		0·59	
A–I	0·65		...		...		...		...		...		0·68		...		0·82		0·71		...	

suggestions such as these, assuming the presence of a detached cloud at about $3r_K$ distance from the limb of the gK5 component.

The question arose whether there would be a *general* absorption progressing at the same time as the *selective* line absorption. The observational proof of this should be easier to obtain in the ultra-violet beyond the head of the Balmer series. For this purpose a group of nine wave-lengths (ranging from 3672 and 3410 A.) has been selected for the measurement of the continuous spectrum of ζ Auri. The problem, from a spectrophotometric standpoint, was a very difficult one, because BD $+41^\circ$ 1044 is far too faint at these wave-lengths, and η Auri, also present in the telescope field, was nearly three magnitudes brighter than the variable. Furthermore, according to some authors, this star may itself be a variable.

As a general rule, a direct comparison of the corresponding spectra of ζ Auri and η Auri was practically impossible. Consequently, the problem was approached in two ways:

(1) The most intense spectrogram of ζ Auri and the weakest one of η Auri have been compared at the same wave-lengths, applying a correcting factor which represents the difference in log-exposure between the two photographs on each plate. This method is not very suitable when spectra with widely different times of exposure are not available on the same plate. Furthermore, the continuous ultra-violet spectrum of η Auri is always much more intense than that of ζ Auri, particularly at the longest wave-lengths. Consequently, it would be necessary to have an absolutely correct standardization curve, covering a wide range of transparencies.

(2) In the spectrum of η Auri a wave-length has been selected far enough in the ultra-violet ($\lambda = 3410$), where densities are easily measurable, and the corresponding intensities have been compared with those deduced for ζ Auri from the simultaneous exposure at the different wave-lengths. Here it is assumed that the absorption due to the terrestrial atmosphere does not change appreciably from 3410 to 3672 A. This second method has the advantage that the transparency of the η Auri spectrum image at 3410 A. is of the same order of magnitude as that of ζ Auri at the various wave-lengths, and too great densities are avoided.

Both methods lead to substantially the same results, as Tables III *a* and III *b* show. We note a good agreement when the resulting differences in $\log I$ at λ 3410 A. are directly compared. It seems very likely that some ten days before the visual eclipse a definite weakening of the continuous ultra-violet spectrum of ζ Auri took place. It might also be mentioned that plate 2, where the lines were particularly intense, shows a lower brightness compared with the subsequent plate.

Huffer (**7**) was the first to suggest, on the occasion of the 1934 eclipse, that the atmosphere of the red component could give rise to a slight drop in the brightness of ζ Auri three or four days before or after the minimum. Miss Swope (**8**) examined about 3000 plates of the Harvard collection and concluded that there was apparently an actual loss of light just before and after the eclipse, amounting to $0^m.10$ and lasting two or three days. She suggests that it is not improbable that the star requires several weeks to reach the maximum magnitude, owing to the extensive atmosphere of the red component. At the time these observations were interpreted as an intrinsic fluctuation of the red component. In our plates such a weakening seems to begin at least eight to nine days before the visual eclipse.

Probably it begins earlier the shorter the wave-length selected for the spectrophotometric measurements.

Whether the comparison star η Auri is variable or not, cannot be ascertained from our plates. It would be desirable to decide this question conclusively with very accurate photoelectric measurements. Anyway, an artificial weakening of this star appears to be necessary if it is to be used as a comparison star for systematic ultra-violet measurements, which are strongly recommended on the occasion of future eclipses.

Finally the present observations show a quicker succession of the various phenomena during the second part of the eclipse, when the B-component is emerging from the atmosphere of the gK5 star. This confirms the previous results (**9**) that here the atmosphere appears to be less extended than in the first part of the eclipse.

The writer wishes to express his gratitude to Professor Redman and Dr Beer for their valuable assistance and interest in this work.

Cambridge :
1949 *March* 31.

References

(1) C. E. Moore, "A Multiplet Table of Astrophysical Interest", Revised Edition, *Princeton Contr.*, **20,** 1945.
(2) A. Beer, *M.N.*, **95,** 24, 1935; **100,** 693, 1940.
(3) W. H. Christie and O. C. Wilson, *Ap. J.*, **81,** 426, 1935.
(4) W. H. Christie, *Publ. A.S. Pac.*, **50,** 53, 1938.
(5) M. G. Fracastoro, *Oss. Mem. Arcetri*, No. 63, 1946.
(6) Z. Kopal, *Ap. J.*, **103,** 310, 1946.
(7) C. M. Huffer, *Ap. J.*, **81,** 292, 1935.
(8) H. H. Swope, *Harv. Circ.*, No. 399, 1935.
(9) D. B. McLaughlin, *Ap. J.*, **108,** 237, 1948.

Note added in proof.—When I wrote this paper, I had assumed that the light of the gK5-component was negligible in the ultra-violet at the wave-lengths which my measurements deal with. Such an assumption was based on the amplitudes published and the very low temperature assumed by many authors for the gK5-component of ζ Auri.

However, further independent measurements of monochromatic amplitudes showed that the effective temperature of the gK5-component is about 3000 deg. K.; so that the alternative interpretation arises that some of the lines in question are merely due to the overlapping of the two spectra. Ionized lines such as *Ti* II λ 3760, appearing about ten days before the visual eclipse, are of atmospheric origin instead.

The general absorption is fully confirmed by the further measurements and agrees with the theoretical predictions.

Harvard College Observatory,
Cambridge,
Massachusetts :
1949 *October.*

FURTHER MEASURES OF THE DECREASE OF WAVE-LENGTH OF SOLAR LINES

J. Evershed, F.R.S.

(Received 1949 August 18)

Summary

Further measures have been made of the absolute wave-lengths of solar lines in spectral regions where they can be compared directly with telluric lines. It is confirmed that in all the lines measured a decrease of wave-length has occurred.

In continuation of my measures of four solar iron lines in the α group of telluric oxygen lines published in *M.N.*, **108,** 347, I have now to record some further measures of this same group, but including a nickel and a strong iron line near the red end of the group. In addition to this I have measured six lines in the region near the D lines, the nickel line between D_1 and D_2 and five iron lines on the side towards red of D_1. For these I use the water-vapour lines of this region as standards of reference for finding the wave-lengths of the solar lines, after allowing for the orbital and diurnal motions of the Earth. In all I have measured three series of lines, and the results are set out in the table following.

Decrease of Wave-length of Solar Lines

Series 1

Element and Int.	*Fe* 5	*Fe* 7	*Fe* 5	*Ni* 4	*Fe* 6
λ	6297	6301	6302	6314	6318
Kodaikanal 1915	·799	·508	·499	·667	·026
Ewhurst 1949 (30)	·795	·501	·496	·661	·020
Decrease in 34 years	·004	·007	·003	·006	·006

Series 2

Element and Int.	*Ni* 4	*Fe* 4	*Fe* 3
λ	5892	5905	5916
Revision of Rowland	·885	·684	·262
Ewhurst 1938 (4)	·886	·682	·258
Ewhurst 1949 (27)	·883	·680	·257
Decrease	·002	·004	·005

Series 3

Element and Int.	*Fe* 3	*Fe* 2	*Fe* 6	*Fe* 5
λ	5916	5927	5930	5934
I.A.U. " Recommended "	·257	·797	·191	·665
Ewhurst 1949 (24)	·254	·793	·188	·662
Decrease	·003	·004	·003	·003

In *Series* 1 the values given in the first row of fractional wave-lengths are from recent measures, at the centre of the Sun, of a negative photographed in the year 1915 and they agree within 0·001 A. with the "recommended" values of the

International Astronomical Union, except for the nickel line 6314 which is not included in the I.A.U. list. The mean values yielded by the 30 Ewhurst spectra of the first three iron lines agree essentially with the 1947 and 1948 spectra as recorded in my previous paper and yield similar decreases. The two lines at the red end of the spectra depend on two oxygen lines assumed at 6314·235 and 6314·974. They also give large decreases.

In *Series* 2, I have measured the nickel line between D_1 and D_2 and two iron lines towards the red end of the spectrum. For these I take the wave-lengths given in the Revision of Rowland and compare them with measures of Ewhurst spectra photographed in 1938 and in 1949. The values obtained depend on seven water-vapour lines interspersed among them, all but one at 5893·513 have been remeasured and marked *M* in the Revision of Rowland.

In this series a progressive decrease of wave-length of the two iron lines is indicated.

Finally in *Series* 3, I use the "recommended" values of the I.A.U., considering that the relative values between the solar and telluric lines may be more accurate than those in the Revision of Rowland. Here, however, the wave-lengths of the 24 Ewhurst spectra depend on only two water-vapour lines at 5919·644 and 5932·092, but the decrease in this series is of the same order as in the other series.

The individual results of each spectrum measured are very consistent, as is indicated by the probable errors computed for each series but not entered in the tables. When the number of measures averaged is as much as or more than 24, the p.e. is a small figure in the 4th decimal, but where it is the mean of only four spectra, as in the 1938 spectra in Series 2, the p.e. is $\pm 0{\cdot}0006$ for each of the three lines. The mean decrease of wave-length of the eleven lines is 0·0042 A.

It is of interest to compare this result with the interferometer measures by Dr Adam of the absolute wave-lengths at the centre of the Sun of lines in the region 6013 to 6270. In Table III of her paper * there are eleven lines common to the I.A.U. list and all of these show smaller wave-lengths, the difference varying from $-0{\cdot}001$ to $-0{\cdot}006$, the mean difference is $-0{\cdot}0033$ A. Her spectra were photographed in the year 1946.

Measures of solar spectra in future years may show whether the wave-lengths of solar lines undergo cyclical variations, which might occur if it is due to a Doppler effect resulting from the turbulent nature of the chromosphere which is known to vary in depth. But a varying movement radial to the Sun would not affect wave-lengths at the limb, yet according to my measures of limb spectra a decrease of wave-length is indicated there also.†

This apparent instability of the solar lines increases the difficulty of finding a satisfactory interpretation of the solar wave-lengths. The general shift of the solar lines towards red compared with the corresponding lines in terrestrial sources and the close agreement in many cases of the shift due to the gravitational effect of Einstein seems to leave little doubt that this is the main cause of the wave-length differences, modified perhaps by variable Doppler effects and by the Lindholm shift discussed by Dr Adam in her paper.

Highbroom,
Ewhurst, Surrey:
1949 *August* 17.

* M. G. Adam, *M.N.*, **108**, 454, 1948.
† *Ibid.*, **108**, 351, 1948.

ON KOWALSKY'S FORMULAE FOR BINARY ORBITS

D. Belorizky

(Communicated by the Foreign Secretary)

(Received 1949 July 12)

Summary

Kowalsky's formulae are derived by a new method.

Introduction.—In 1873 Kowalsky gave an analytical method of derivation of the geometrical elements of the real orbit of visual binary stars, the apparent orbit being known. This method has been reproduced by T. J. J. See (**1**), Ch. André (**2**) and R. G. Aitken (**3**) in their respective treatises.

Let

$$Ax^2+2Bxy+Cy^2+2Dx+2Ey-1=0 \qquad \text{(I)}$$

be the equation of an apparent binary orbit, the principal star being at the origin of the rectangular axes Ox, Oy in the plane P of the apparent orbit. Ox lies in the direction of the North point, Oy eastward. Kowalsky resolves the following problem:

Given equation (I), find the geometrical elements of the real orbit, i.e. elements Ω, i, e, ω and a, where Ω is the angle between the line of nodes and Ox, ω the angle between the semi-major axis and the line of nodes, i the inclination of the orbit, e its eccentricity and a its semi-major axis.

Kowalsky writes these elements as functions of the coefficients in (I); however, his method of derivation is lengthy and cumbersome.

W. M. Smart (**4, 5**) undertook a simplification in 1930, but his derivation is still somewhat complicated.

In the present paper Kowalsky's formulae are derived from quite different considerations, believed to be much simpler.

1. Consider the real orbit. The principal star is at one of the foci, O_1, projected in O. In the plane P_1 of the real orbit take axes O_1u, v determined as follows: O_1u is parallel to the line of nodes, i.e. to the intersection of planes P_1 and P; the positive direction is taken from O_1 to the periastron. The axis O_1v is at 90° with O_1u in the same sense as Oy relative to Ox in the apparent orbit.

Take O_1u as polar axis; let W be the angle between the radius vector ρ and O_1u: the equation of the orbit in polar coordinates is

$$\rho=\frac{p}{1+e\cos(W-\omega)}, \qquad (1)$$

where p is the parameter.

Equation (1) yields $\rho+\rho e(\cos W\cos\omega+\sin W\sin\omega)=p$. But $\rho\cos W=u$ and $\rho\sin W=v$; thus

$$\frac{\rho^2}{p^2}=\frac{u^2+v^2}{p^2}=\left(1-\frac{e}{p}u\cos\omega-\frac{e}{p}v\sin\omega\right)^2. \tag{2}$$

Project now the real orbit as well as axes O_1u, O_1v on plane P of the apparent orbit: let Ox' and Oy' be the projections of the axes. Ox' is parallel to the line of nodes, therefore

$$x'=u, \qquad y'/\cos i=v, \tag{3}$$

and replacing in (2) u and v by their expressions (3),

$$\frac{1}{p^2}\left(x'^2+\frac{y'^2}{\cos^2 i}\right)=\left(1-\frac{e}{p}x'\cos\omega-\frac{e}{p}y'\frac{\sin\omega}{\cos i}\right)^2. \tag{4}$$

Equation (4) evidently represents conic section (I) with axes $Ox'y'$.

2. Now Ox' makes an angle Ω with Ox of (I). Rotate axes $Ox'y'$ by angle $-\Omega$, in order to make them coincide with Oxy:

$$\begin{cases} x'=x\cos\Omega+y\sin\Omega, \\ y'=-x\sin\Omega+y\cos\Omega; \end{cases}$$

(4) then becomes

$$\frac{x^2}{p^2}(1+\sin^2\Omega\tan^2 i)-\frac{2xy}{p^2}\cos\Omega\sin\Omega\tan^2 i+\frac{y^2}{p^2}(1+\cos^2\Omega\tan^2 i)$$

$$=\left[1-\frac{e}{p}x\left(\cos\Omega\cos\omega-\sin\Omega\frac{\sin\omega}{\cos i}\right)-\frac{e}{p}y\left(\sin\Omega\cos\omega+\cos\Omega\frac{\sin\omega}{\cos i}\right)\right]^2. \tag{5}$$

For the sake of brevity, write (5) as

$$\alpha x^2+2\beta xy+\gamma y^2=(1-\delta x-\epsilon y)^2, \tag{6}$$

α, β, γ, δ, ϵ having evident significations.

Equations (5) or (6) must be identical to (I), as representing the same curve; their independent coefficients are the same, viz. -1. But in the form (6), the equation of the apparent orbit is that of a conic section tangential to the two straight lines passing through the origin and having the equations

$$\alpha x^2+2\beta xy+\gamma y^2=0;$$

the points of contact are those of intersection with the straight line $1-\delta x-\epsilon y=0$. This line is therefore the polar of the origin.

Moreover, (2) shows that the equation

$$1-\frac{e}{p}u\cos\omega-\frac{e}{p}v\sin\omega=0$$

is that of the directrix of the real orbit. Now the directrix is the polar of the focus; thus the projection of the polar of focus O_1 becomes the polar of the origin O.

Therefore $1-\delta x-\epsilon y=0$ is the equation of the polar of the origin with respect to the apparent orbit.

Equation (I) will now be put in form (6). The equation of the polar of the point (x_0, y_0) with respect to (I) is

$$x(Ax_0 + By_0 + D) + y(Bx_0 + Cy_0 + E) + Dx_0 + Ey_0 - 1 = 0.$$

At the origin, $x_0 = y_0 = 0$: the equation of the polar of the origin is therefore

$$Dx + Ey - 1 = 0.$$

But

$$(Dx + Ey - 1)^2 = D^2x^2 + E^2y^2 + 2DExy - 2Dx - 2Ey + 1,$$

and taking (I) into account,

$$(1 - Dx - Ey)^2 = (A + D^2)x^2 + 2(B + DE)xy + (C + E^2)y^2 \qquad \text{(II)}$$

identical to (6): therefore

$$\frac{1}{p^2}(1 + \sin^2 \Omega \tan^2 i) \equiv \alpha \equiv A + D^2, \qquad (7)$$

$$-\frac{1}{p^2}\sin \Omega \cos \Omega \tan^2 i \equiv \beta \equiv B + DE, \qquad (8)$$

$$\frac{1}{p^2}(1 + \cos^2 \Omega \tan^2 i) \equiv \gamma \equiv C + E^2, \qquad (9)$$

$$\frac{e}{p}\left(\cos \Omega \cos \omega - \sin \Omega \frac{\sin \omega}{\cos i}\right) \equiv \delta \equiv D, \qquad (10)$$

$$\frac{e}{p}\left(\sin \Omega \cos \omega + \cos \Omega \frac{\sin \omega}{\cos i}\right) \equiv \epsilon \equiv E. \qquad (11)$$

Adding and subtracting (7) and (9),

$$\frac{1}{p^2}(2 + \tan^2 i) \equiv \gamma + \alpha \equiv C + E^2 + A + D^2, \qquad \text{(III)}$$

$$\frac{1}{p^2}\cos^2 \Omega \tan^2 i \equiv \gamma - \alpha \equiv C + E^2 - A - D^2. \qquad \text{(IV)}$$

Equation (8) can be written as

$$\frac{1}{p^2}\sin 2\,\Omega \tan^2 i \equiv -2\beta \equiv -2(B + DE). \qquad \text{(V)}$$

Lastly, (10) and (11) can be replaced by

$$\frac{e}{p}\cos \omega = D \cos \Omega + E \sin \Omega, \qquad \text{(VI)}$$

$$\frac{e}{p}\sin \omega = (-D \sin \Omega + E \cos \Omega)\cos i. \qquad \text{(VII)}$$

Equations (III) to (VII) represent Kowalsky's formulae as given in the astronomical treatises.

3. It may be worth while to add two formulae which may serve as a check when the elements of the real orbit have been calculated by means of (III) to (VII).

The expression $(A+D^2)(C+E^2)-(B+DE)^2$ is equal to $-\Delta$, where Δ is the discriminant of (I). Now from (7), (8) and (9)

$$-\Delta=-\begin{vmatrix} A & B & D \\ B & C & E \\ D & E & -1 \end{vmatrix}=(A+D^2)(C+E^2)-(B+DE)^2=\frac{1}{p^4\cos^2 i}. \quad \text{(VIII)}$$

On the other hand, putting

$$\delta_1=AC-B^2,$$

$$-\frac{\Delta}{\delta_1}=\frac{1}{1-e^2}. \quad \text{(IX)}$$

This formula can be derived by direct calculation of A, B, C in terms of the elements, comparing equations (5) and (I). But it has been derived from simpler considerations in another paper (**6**).

The Observatory,
Marseilles:
1949 *June* 3.

References

(1) T. J. J. See, *Researches on the evolution of the stellar systems*, Vol. I, pp. 45–49, 1896.
(2) Ch. André, *Traité d'Astronomie stellaire*, Vol. II, pp. 26–31, 1900.
(3) R. G. Aitken, *The Binary Stars*, pp. 74–78, 1918.
(4) W. M. Smart, " On the derivation of the elements of a visual binary orbit by Kowalsky's method ", *M.N.*, **90,** 534–538, 1930.
(5) See also R. G. Aitken, *The Binary Stars*, second edition, pp. 81–85, 1935.
(6) D. Belorizky, " La loi de Newton, le mouvement des étoiles doubles et le problème de Bertrand ", *Journal des Observateurs*, **30,** 192, 1948.

THE ORIGIN AND EVOLUTION OF THE SOLAR SYSTEM

K. E. Edgeworth

(Received 1949 June 7)

Summary

Various theories are reviewed. By eliminating ideas which appear to be faulty, by accepting ideas which appear to be sound and helpful, and by introducing some new ideas to fill the gaps, the outline of a complete theory is built up.

The only mechanism that appears to be capable of explaining the regularities of the solar system is fluid friction, and it is inferred that the material which now forms the planets and their satellites existed at one time in the form of a vast rotating disk. Bodies of planetary mass cannot be formed by condensation in a cloud of gas, and small solid particles must have been the chief ingredient in this primitive material. The angular momentum of the disk is explained quite simply by the assumption that the material was collected from interstellar space, and this is the only explanation that gives a result which is numerically of the right order.

Condensations would form in the rotating disk, and in the first instance these would be small and numerous. Subsequently they would coalesce to form a single large cluster in each region of interplanetary space. Each cluster would consist of a nucleus and a surrounding disk or annulus. Owing to the combined effect of viscosity and tidal friction, there would be a loss of angular momentum and the bulk of the material would be absorbed by the nucleus to form the planet. The residue of material in the annulus would condense to form the satellites.

In the region which lies immediately inside Jupiter's orbit the perturbations caused by the planet would drive much of the material inwards towards the Sun. In the attenuated residue the condensations would be smaller and more numerous than usual; they would coalesce to some extent, but not sufficiently to form a single planet, and the result would be the formation of the group of small bodies we know as the asteroids. In the region outside the orbit of Neptune the material would also be highly attenuated, and here again condensations would be small and numerous, but the progress of evolution was slower, and the region is probably populated by a very large number of small clusters. Wandering clusters make their appearance from time to time as comets.

Introduction

1. The paper is based on the assumption that there is sufficient knowledge available to permit of the construction of a complete theory of the origin and evolution of the solar system, provided that the information is carefully sifted and properly utilized. Helpful ideas must be retained, faulty ideas must be eliminated and gaps in the resulting theory must be carefully filled.

2. A complete theory necessarily raises a number of questions of a controversial character and a full discussion would need a book rather than a paper. Moreover, the various digressions that would be involved in this method of approach would tend to obscure the main argument. In order to present a clear picture of the whole process, therefore, detailed discussion of particular issues will be avoided, on the understanding that the argument can be expanded on some future occasion if required.

Part I. *The regularities of the system and their significance*

3. The planets revolve about the Sun in the same direction and in nearly the same plane, and these and other regularities of the system are much too remarkable to be due to chance. They must be ascribed to the existence of some common cause (Laplace).*

4. The only mechanism that is known to be capable of creating the required regularity is fluid friction, the term fluid being taken to include a cloud of small solid particles as well as a cloud of gas. The inference is that the material that now forms the planets and their satellites existed at one time in the form of "a fluid of immense extent" (Laplace). The alternative hypothesis that the regularities were produced after the appearance of the embryo planets is rejected for reasons which are set out below.

5. The bulk of the materials of which the Earth is composed are such that they would necessarily assume the solid state at the temperature of interplanetary space, and this statement is no doubt true also of the other terrestrial planets. There is the further point that condensations of planetary mass cannot be formed in a gas.† For these reasons it may be inferred that the material out of which the planets were originally formed consisted of small solid particles.

6. Although a cloud of small solid particles was the essential ingredient out of which the planets were first formed, gas must also have been present. When the major planets reached a certain size they captured gas as well as other materials.

7. The angular momentum of the planets must have been derived from one or more of three sources: from the Sun, from some external source, or from the original material.

8. Viscous friction would provide a mechanism capable of transferring energy and angular momentum from the Sun to a surrounding cloud of fluid, but calculation shows that it could not be responsible for the transfer of an adequate amount in the available time.

A theory was put forward a short time ago by Alfvén in which it is suggested that electromagnetic forces might take the place of viscous friction‡, but the possibility that such a process might be significant remains to be demonstrated.

9. The theory, which has sometimes been put forward, that a planet might be born ready-made from the Sun (or a satellite from a planet) is untenable. Firstly, because a body ejected from the Sun would not possess sufficient angular momentum to establish it in a planetary orbit; secondly, because an explosion capable of ejecting the material would not eject large masses, but would scatter the material in all directions.

10. The nebular theory usually associated with the name of Laplace postulates that the Sun and the planets were condensed from the same rotating cloud of gas. The theory is untenable because the angular momentum per unit mass of the planets differs much too widely from the angular momentum per unit mass of the Sun.§ The inference which must be drawn from the facts is that the formation of the Sun and the formation of the planets were two distinct events which occurred at different times and were based on different sources of material.

* *Exposition du Système du Monde*, Paris.

† K. E. Edgeworth, *M.N.*, **106**, 473, 1946.

‡ A. Hunter, *Nature*, **156**, 378, 1945.

§ The first writer to direct attention to the importance of angular momentum in relation to theories of planetary evolution was M. Babinet, *Comptes Rendus*, **52**, 481, 1861.

11. Collision theories represent an attempt to explain the dual origin of the material. They are open to the following objections*:—

(*a*) They lack plausibility. The number of collisions of the assumed type that have taken place since the formation of the stars is probably only a few hundreds.

(*b*) The filament which is supposed to have been formed as a result of the collision would not break up into ready-made planets but would disintegrate.

(*c*) The angular momentum of the planets (or material) resulting from the collision would be inadequate.

(*d*) The orbits of the planets which would be formed by this process would be highly eccentric.

(*e*) The angular momenta of the satellites about their primaries create difficulties of precisely the same kind as do the angular momenta of the planets about the Sun, but it is impossible to invoke special collisions for each of the planets. "The systems of Saturn and Jupiter are so like that of the Sun that any hypothesis which assigned different origins to the system and its sub-systems would be condemned by its own artificiality" (Jeans).†

12. The theory of a separate collision between two other stars, one of which was a companion to the Sun, appears to meet the objection that the angular momentum is insufficient, but it does not meet any of the other objections.

13. It has been suggested that the rounding up of eccentric orbits might have been brought about by the presence of a "resisting medium", and the influence of this resisting medium has been compared to the influence exerted by the air in a clock-case on the oscillations of a pendulum. The analogy is false. The molecules of air in a clock-case are not absorbed by the pendulum and the air maintains its separate existence indefinitely. Small particles approaching a planet, on the other hand, would be absorbed. To be of any significance the mass of the resisting medium would have to be of the same order as the mass of the planets.

The questions then arise: what was the resisting medium? whence did it come? and what has happened to it? No serious attempts to answer these questions appear to have been made, and, if an attempt is made to answer them, we seem to be driven back inevitably to our original proposition that "the Sun was at one time surrounded by a fluid of immense extent".

14. Other solutions having failed, we return to the assumption that the angular momentum of the planets was contained in the original material, and we postulate a vast rotating disk made up of separate particles each revolving about the Sun in its own orbit. To emphasize this aspect of the matter the individual particles may be referred to as "planetesimals" (Chamberlin and Moulton).‡ The rotating disk would resemble the rings of Saturn but on a much vaster scale.

15. Collisions between the particles are evidently of importance. At first they might be of considerable violence and the individual particles might be vaporized, but they would condense again.§ As time went on, random velocities would be reduced and collisions between particles would become elastic but not completely so; further reduction in the random velocities would occur and the disk would become very thin. We are again reminded of the similarity between this postulated disk of planetary material and Saturn's rings.

* See Harold Jeffreys, *M.N.*, **108**, 94, 1948.

† J. H. Jeans, *Problems of Cosmogony and Stellar Dynamics*, Cambridge, p. 281, 1919.

‡ T. C. Chamberlin, *Ap. J.*, **14**, 17–40, 1901. Chamberlin and Salisbury, *Geology: Earth History*, Vol. II, 38–81, 1906.

§ A. L. Parson, *M.N.*, **105**, 244, 1945.

16. There is a very simple and obvious theory which can be offered to explain the origin of the planetary material, and it is surprising that it seems never to have been seriously considered.

It would be unreasonable to suppose that the formation of the stars involved the capture of the whole of the scattered material in interstellar space. At least a small residue would remain uncaptured, and this residue would be supplemented by matter ejected by stellar explosions. A star such as the Sun would eventually capture the stray material in its own neighbourhood, and this captured material would condense to form the vast rotating disk which has already been described.

Similarly it may be postulated that the process, whatever it was, that led to the formation of the planets did not absorb the whole of the available material. A residue remained which subsequently condensed to form the satellites.

17. The theory that the planetary material was condensed from interstellar space may be tested as follows. The volume of space occupied at the present time by a star such as the Sun is about 800 cubic light-years. At the time when the stars were formed it was probably about one hundredth part of this *, say 8 cubic light-years $= 6{\cdot}7 \times 10^{54}$ cub. cm., equivalent to a sphere of radius $1{\cdot}2 \times 10^{18}$ cm. The orbital angular velocity of the Sun about the centre of the Galaxy is 10^{-15}, and the local angular velocity of a cloud of material would be about one-third of this. The surface angular momentum per unit mass of the supposed cloud of material in c.g.s. units is therefore 5×10^{20}, compared with $2{\cdot}5 \times 10^{20}$ which is the angular momentum per unit mass of the planet Neptune. It will thus be seen that the theory gives a value of the planetary angular momentum which is exactly of the right order, and it is the only theory which does so.

18. A small part of the material of the rotating disk was absorbed by the Sun, and this explains why the rotation of the Sun is in the same direction as the motion of the planets in their orbits.

19. The preceding sections appear to establish, beyond all reasonable doubt, that the planets were evolved from a vast rotating disk of scattered material, and that this material was originally derived by capture from interstellar space.

It remains to consider by what processes this rotating disk of scattered material condensed to form the planets and their satellites.

Part II. *The formation of the planets and their satellites*

20. This part of the paper starts with the postulate that the material out of which the embryo planets were first formed existed originally in the form of a rotating disk of small solid particles extending beyond the orbit of Neptune.

21. Any particular part of the disk may be stable or there may be a tendency to form local condensations.

Let R = distance from the centre of the primary,
M = mass of the system inside the radius R,
$\bar{\rho}$ = mean density of the system $= 3M/4\pi R^3$,
A = radius of the local region,
m' = mass of the material inside the radius A, and
ρ' = local density $= 3m'/4\pi A^3$.

* K. E. Edgeworth, *M.N.*, **106**, 472, 1946.

Jeffreys * has shown that a local condensation cannot exist unless

$$\rho' > 0{\cdot}022\bar{\rho}. \qquad (1)$$

Maxwell's formula for the rings of Saturn † can be made to yield a similar result, that is to say

$$\rho' > 0{\cdot}038\bar{\rho}. \qquad (2)$$

Both formulae are open to the criticism that they involve an over-simplification of the problem. A more complete analysis would no doubt yield a different value of the numerical coefficient, although it could hardly alter the general character of the formula. Whatever the actual formula may be, if ρ' is less than the critical value, the disk is stable and the formation of local condensations is impossible. On the other hand, if the critical value is exceeded, condensations will be formed and will grow at the expense of the rest of the disk.

22. As a result of collisions between the particles, random velocities are soon reduced to a low figure and the disk becomes very thin. It will then be found that the condition defined by equation (1), or whatever the accepted formula may be, is satisfied, provided that the size of the condensation does not exceed a certain limit.

Let it be assumed that the area of the region from which each planet derived its material was $2R^2$.

Let N = number of condensations formed in this region, and m = mass of the planet.

Then it follows from equation (1) that

$$N = m/m' > 10^{-4} \times (M/m)^2. \qquad (3)$$

Applying this formula to the case of Jupiter, the number of condensations needed to form the planet is found to be several hundreds; in the case of the Earth they would be numbered in millions. A different value of the numerical coefficient in the previous equation would involve a corresponding change in equation (3). In fact, however, the exact number of condensations formed is unimportant; the essential point is that they would be numerous and small in size.

It may be remarked in passing that the estimate of the number of condensations formed does not enter into our subsequent discussion as to the time required for the condensations to coalesce to form a planet, for the reason that the time required for the early stages is short in comparison with the total time.

23. The condensations that might be formed in the rotating disk would, in the first instance, be in the nature of eddies, that is to say they would maintain a definite structure, but the "fluid" would flow through them. They would be regions in which the local density was above the average.

Flow within the eddy would be subject to viscous friction, and the effect of the friction would be to cause the eddy to grow at the expense of the rest of the disk.

24. Eventually the central density of an eddy would increase sufficiently to make it possible for particles to move in closed orbits about the centre. This is equivalent to the condition that ‡

$$\rho' > 3\bar{\rho}.$$

The condensations may now be described as "clusters", and the central part of a cluster may be described as the "nucleus".

25. At first the clusters would consist simply of regions of high density embedded in a continuous sheet of material. With lapse of time, viscosity would cause the clusters to grow by absorbing material from the disk until the whole of

* Harold Jeffreys, *The Earth*, Cambridge, p. 15, 1924.
† K. E. Edgeworth, *M.N.*, **106**, 481, 1946.
‡ K. E. Edgeworth, *loc. cit.*, p. 481.

the scattered material was absorbed, and the system would then consist of a number of individual clusters separated from one another by regions of empty space.

26. It is at this point that erroneous inferences appear to have been drawn as to the course of events. It has been assumed that the clusters would contract upon themselves, and that the final result would be, not a small number of large planets, but rather a large number of small planets such as the asteroids. It has therefore been argued that the planets could not have been evolved by a process of condensation from a rotating disk of scattered material, and other solutions have been sought. The argument is not valid.

27. There were in fact two evolutionary processes at work. The clusters did contract upon themselves, but at the same time they tended to coalesce, so that the average mass of a cluster increased. The course of events was determined by the relative importance of these two processes.

28. Following upon the absorption of the scattered material by the clusters, the size of a cluster is determined by its rotation, and contraction depends upon the dissipation of angular momentum, or more precisely on the transfer of angular momentum from the individual cluster to the system as a whole. The mechanism responsible for the loss of angular momentum is akin to tidal friction and may be described as such.

With tidal friction the magnitude of the effect declines as the radius of the cluster decreases, so that contraction proceeds at an ever decreasing rate.

29. The actual rate of contraction may be estimated as follows. When first formed, the boundary of a cluster may be identified with the boundary within which particles can move in closed orbits. This region is ellipsoidal, but for simplicity it may be regarded as spherical.

Let A = the radius of a cluster (as defined above), then

$$A^3 = R^3(m'/3M).* \tag{4}$$

It is clear that inside the boundary the gravitational field is mainly due to the cluster itself; outside the boundary the field of the Sun is of dominating importance. It may therefore be said that the boundary itself defines a surface on which the attractions of the Sun and the cluster are of equal importance. Let it now be assumed that, under these conditions, the radius of the cluster is reduced by one per cent in each revolution; the precise figure is not of serious importance to our present argument, which would not be materially affected if our assumption were altered to 1 in 10 or 1 in 1000.

Let a = the radius of the cluster after time t.

It has been shown by Darwin † that the retardation due to tidal friction is inversely proportional to the sixth power of the distance, and it follows that it is also proportional to the sixth power of the diameter, since it is the ratio between diameter and distance that is involved in the argument. Combining these assumptions we get

$$\frac{dt}{da} = -\frac{100}{A}(A/a)^6 p, \tag{5}$$

where p is the period of revolution at distance a. Now p is itself a function of a, that is to say

$$p = (a/A)^{\frac{3}{2}} P', \tag{6}$$

---

* *Loc. cit.*, p. 481.

† G. H. Darwin, *The Tides*, London, p. 247, 1898.

where P' is the period of revolution at distance A. Remembering that the local angular velocity of the condensation is about one-half the angular velocity of the condensation in its orbit, we may write $\frac{1}{2}P' = P =$ the period of revolution about the Sun, so that

$$dt/da = -200\alpha^{-\frac{1}{2}}A^{\frac{1}{2}}P, \tag{7}$$

$$t = 57(A/a)^{\frac{1}{2}}P, \tag{8}$$

$$a/A = 3(t/P)^{-\frac{2}{3}}. \tag{9}$$

This equation gives the rate at which a cluster contracts upon itself as a result of tidal friction due to the action of the Sun.

30. We must now turn to the question of stability in a system which consists of a disk made up of small clusters, each cluster being itself a disk of small solid particles. If Maxwell's analysis is applied to this type of structure, it is found to be stable, and any oscillations that may occur therein are of constant amplitude. It is necessary, however, to take into account certain factors which are not discussed in Maxwell's essay.

Until such time as scattered material between the clusters is absorbed, relative motion of adjacent clusters is subject to resistance due to viscous friction. If allowance is made for this damping, it would no doubt be found that the amplitude of the oscillations is no longer constant.

Clusters at different distances from the Sun move in their orbits with different angular velocities and the distance of a cluster from its neighbours is not constant, even if it is postulated that oscillations are originally absent.

In the long run close approaches between clusters lead to an increase in the random velocities.*

All these factors tend to increase the relative motion of adjacent clusters, so that collisions take place between one cluster and another. If the collision is slight the clusters may separate again, but, in the case of a head-on collision, the colliding clusters must coalesce. It will be assumed that in one in every five collisions the colliding clusters fail to separate again and coalesce to form a single cluster.

31. The rate at which clusters collide and coalesce depends on the random velocity and on the size of the clusters. For the purpose of numerical calculation the random velocity is assumed to be one hundredth part of the velocity of the cluster in its orbit. When two clusters coalesce the mass of the new cluster is the sum of the original masses, and it is assumed that equations (4) and (9) are still applicable and can be used to determine its radius. The rate at which the number of clusters is reduced by collisions may then be estimated as follows.

In our present problem the frequency of the collisions must be regarded as a problem in two dimensions rather than three, and the duration of the free path is of the order

$$\frac{0{\cdot}2}{vn(2a)},$$

where $v =$ random velocity $= 0{\cdot}01R\omega = 0{\cdot}06R/P$,

$n =$ number of clusters per unit area $= N/2R^2$, and

$a =$ radius of the cluster $= 3R(m/3MN)^{\frac{1}{3}}t^{-\frac{1}{3}}P^{\frac{1}{3}}$,

so that the duration of a free path becomes

$$1{\cdot}6(M/m)^{\frac{1}{3}}N^{-\frac{1}{3}}t^{\frac{1}{3}}P^{\frac{1}{3}}. \tag{10}$$

The number of collisions per unit time is the reciprocal of this multiplied by $\frac{1}{2}N$.

* K. E. Edgeworth, *M.N.*, **108**, 417, 1948.

It is assumed that one collision in five causes the colliding clusters to coalesce, so that

$$\delta N/\delta t = -0{\cdot}06(m/M)^{\frac{1}{3}}N^{\frac{5}{3}}t^{-\frac{3}{7}}P^{-\frac{5}{7}}. \qquad (11)$$

A solution to this equation is given by

$$t/P = 400(M/m)^{7/15}N^{-14/15}. \qquad (12)$$

To obtain the time required to reduce the number of clusters from (say) 10^6 to one, we substitute these values for N in equation (12) and we get

$$(t_2 - t_1)/P = 400(M/m)^{7/15}(1 - 0{\cdot}0000025) = 400(M/m)^{7/15}.$$

For the Earth	$t/P = 1{\cdot}5 \times 10^5$,	$t = 1{\cdot}5 \times 10^5$ years
Jupiter	$t/P = 10^4$,	$t = 1{\cdot}2 \times 10^5$,,
Neptune	$t/P = 4 \times 10^4$,	$t = 6 \times 10^6$,,

It is therefore clear that the clusters in any particular region will coalesce to form a single large cluster within a comparatively short time.

32. Our next problem is to explain how it came about that the vast rotating cluster of material got rid of its angular momentum and condensed to form a planet.

Throughout this phase the cluster consisted of two parts: the central condensed portion, or nucleus, and a rotating disk which may be referred to as the annulus. Broadly speaking, the explanation is that angular momentum was transferred outwards from the nucleus to the periphery of the annulus by viscous friction, and that, having reached the periphery, it was dissipated by tidal friction as already described.

33. At this stage in our analysis it is convenient to introduce the idea of opacity, which is defined as follows:—

Let n = the number of particles per unit area of the disk,
b = the radius of a particle,
q = the mass of a particle, and
σ = the mass of the disk per unit area.

The proportion of light that is intercepted by a layer of widely scattered material is πnb^2, and this quantity will be called the opacity, irrespective of whether the disk is able to transmit light or not; it may be denoted by Q.

Now $\sigma = nq$, and therefore

$$Q = \pi b^2 \sigma/q. \qquad (13)$$

For the purpose of numerical calculation it may be assumed that $b = 0{\cdot}25$ cm., $q = 0{\cdot}2$ g., which gives approximately

$$Q = \sigma,$$

and on this basis it is possible to prepare a table showing the approximate value of Q for the material which was destined to form the various planets.

TABLE I

Values of $Q = \sigma = m/2R^2$

Mercury	30	Jupiter	160
Venus	20	Saturn	13
Earth	13	Uranus	1
Mars	0·5	Neptune	0·2
Asteroids	0·001 (?)	Pluto	0·001 (?)

It will be observed that the values range from 160 for Jupiter to 0·2 for Neptune; Pluto must be regarded as an escaped satellite.

34. So long as the annulus remained sufficiently opaque, viscosity continued to be significant, and the outflow of angular momentum and the inflow of material continued. When the supply of material fell off, however, the value of Q decreased and viscosity decreased also. Eventually viscosity ceased to be effective, the inflow of material was checked, and the residue remained available for the formation of the satellites.

The approximate mean value of the opacity for the region in which Jupiter's larger satellites were formed, on the assumption that the radius of the region was about $2{\cdot}5 \times 10^{11}$ cm., would be about 10^3. During the earlier period, while the planet itself was being formed, the opacity must have been higher.

35. The mechanism involved in the evolution of the satellites from the rotating disk of material surrounding the planet was no doubt similar to that involved in the evolution of the planets from the rotating disk of material surrounding the Sun, and calls for no special comment. At the same time it may be remarked that the evolution of the planets took place without external interference, with the result that the system exhibits a high degree of regularity, whereas the evolution of the planetary systems was subject to interference by the Sun, so that the smaller systems are less regular than the larger.

36. The inner satellites of Jupiter, Saturn and Uranus move in orbits which are nearly in the same plane as the rotation of the planet itself, from which it may be inferred that there was some close connection between the evolution of the planet and the evolution of the principal satellites. The theory set out in the previous sections explains why this was so.

37. The outer satellites of these three major planets are small and their orbits are retrograde. The small satellites constitute a special problem which will not be discussed here.

38. The evolution of the Earth and Moon also constitutes a special problem, which will not be discussed here.

Part III. *Asteroids and comets*

39. Reference must now be made to the case in which the opacity of the rotating disk is appreciably less than unity. The present mass of the asteroids suggests that the opacity of the original material was of the order 0·001 or thereabouts, but it is possible that an appreciable quantity of the original material has since been captured by some of the planets, so that the initial opacity may have been higher.

40. In a rotating disk, free from external interference, the random velocity of the particles might be so low that collisions would be practically eliminated. On the other hand, the orbits of the particles would be perturbed by the planets and random velocities would be increased. It is difficult to say therefore to what extent viscosity would be developed in a transparent disk in the region now occupied by the asteroids.

If the formation of eddies was possible, we may write as before

$$\rho' > 0{\cdot}022\bar{\rho} \quad \text{and} \quad N > 10^{-4}(M/m)^2.$$

On the other hand, if the boundary of a condensation is identified with the region within which a particle can move in a closed orbit, we have

$$\rho' > 3\bar{\rho} \quad \text{and} \quad MN > 4{\cdot}5(M/m)^2.$$

In the former case N is found to be of the order 10^{18} and in the latter case 10^{13}.

All that can be said with certainty is that the condensations would be much smaller and more numerous than in the case of other planetary material.

41. In accordance with the argument developed in earlier sections of this paper the condensations or clusters would come into collision with one another and would coalesce, so that in the course of time the average mass would increase and the number would be reduced. Owing to the much longer time involved, random velocities would be higher and the problem of collisions must be dealt with in three dimensions, so that the duration of the free path is of the order

$$\frac{0{\cdot}2}{vn(2a)^2},$$

where v = random velocity $= \frac{1}{2}z\omega = \pi z/P$,

z = half the thickness of the disk,

n = number of clusters per unit volume $= N/(4R^2z)$,

a = radius of cluster $= 3R(m/3MN)^{\frac{1}{3}}t^{-\frac{2}{7}}P^{\frac{2}{7}}$;

so that the duration of the free path becomes

$$0{\cdot}007(M/m)^{\frac{3}{7}}N^{-\frac{1}{3}}t^{4/7}P^{\frac{3}{7}}, \tag{14}$$

from which we obtain

$$\delta N/\delta t = -14(m/M)^{\frac{3}{7}}N^{\frac{4}{3}}t^{-4/7}P^{-\frac{3}{7}}, \tag{15}$$

and thence

$$t/P = 0{\cdot}034(M/m)^{14/9}N^{-\frac{7}{3}}. \tag{16}$$

The mass of the asteroids is not known exactly, but M/m is probably of the order 10^9. We may also write $t = 4 \times 10^9$, $P = 4$, and thus the number of asteroids at the present time is found to be

$$N = 1800,$$

which is a result of the right order. A disk of such small mass would not coalesce to form a single planet.

42. It would be unreasonable to suppose that the original rotating disk of scattered material came to an abrupt end outside the orbit of Neptune. There must have been a gradual thinning out of the material at the outer boundary. There is no definite evidence as to the opacity of this material except that it was insufficient to lead to the formation of a planet.

43. Comparing the clusters formed outside the orbit of Neptune with the asteroids, the chief point of difference is that the evolutionary process was much slower. The tempo of the process depends on the factor t/P, and the period of rotation P is about one hundred times greater in the former case than in the latter. Under similar conditions the process of evolution would be much less advanced and the clusters would be smaller and more numerous.

It is not unreasonable to suppose that this outer region is now occupied by a large number of comparatively small clusters, and that it is in fact a vast reservoir of potential comets. From time to time one of these clusters is displaced from its position, enters the inner regions of the solar system, and becomes a visible comet.

Conclusion

The analysis set out in the paper leads to the conclusion that there is no need to invoke some special and improbable coincidence, such as a stellar collision, in order to explain the origin of the solar system. On the contrary, the solar system can properly be regarded as a natural product of normal evolutionary forces.

Cherbury, Booterstown,
Co. Dublin :
1949 *June* 3.

ERRATUM

M.N., **109**, No. 1, R. E. Williamson and G. F. D. Duff's paper :

P. 48, equation (8*a*) for $\log C = -28{\cdot}5283 + 0{\cdot}3276\alpha$ etc.

please read : $\log C = -30{\cdot}7236 + 0{\cdot}3276\alpha$ etc.